OM protein - protein search, using sw model

Run on: September 17, 1999, 03:08:16 ; Search time 64.1 Seconds
             (without alignments)
             11.086 Million cell updates/sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

Title:
Perfect score: 139
Sequence: US-09-037-657-32
Sequence: 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36:*

     Pred. No. is the number of results predicted by chance to have a,
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	139	100.0	166	1	P50161	Sequence encoded b
2	139	100.0	30	1	W10638	Murine IL-3 signal
3	139	100.0	30	1	W00823	Interleukin-3 sign
4	139	100.0	359	1	W56260	Construct containi
5	139	100.0	30	1	W55017	Murine IL3 signal
6	48	34.1	694	1	W31267	Drosophila frizzle
7	46	33.1	116	1	R71911	Anther specific pro
8	45	32.4	96	1	W46836	Bacillus thuringie
9	45	32.4	652	1	W30637	Human 7-transmembr
10	45	32.4	521	1	W30638	Partial human 7-tr
11	43	30.9	1268	1	W58774	Human breast cance
12	43	30.9	481	1	W55029	G-protein coupled
13	43	30.9	542	1	W55030	G-protein coupled
14	42	30.2	266	1	W71460	Protein encoded by
15	42	30.2	306	1	W90236	E. coli antibiotic
16	42	30.2	299	1	W90237	E. coli antibiotic
17	42	30.2	266	1	W90238	E. coli antibiotic
18	42	30.2	51	1	Y02760	Human secreted pro
19	42	30.2	55	1	Y12970	Protein sequnce o
20	42	30.2	428	1	W63713	Human hIK1 protein
21	41.5	29.9	427	1	R75757	Human calcium acti
22	41.5	29.9	315	1	R43568	Mouse MSH-R. Melan
23	41	29.5	709	1	R99801	CRII-7 nerve prote
24	41	29.5	315	1	W37826	Mouse melanocyte s
25	41	29.5	315	1	W87864	Mouse melanocortin
26	41	29.5	336	1	R28149	Sugar beet beta-1.
27	40	28.8	1276	1	R35199	Mouse multidrug re
28	40	28.8	30	1	R63638	Sheep erythropoiet
29	40	28.8	22	1	R65246	N-terminal of 48 k
30	40	28.8	308	1	W98017	BAV3 ORF2 product.
31	40	28.8	523	1	W78915	Bovine butyrophili
32	40	28.8	248	1	W94567	Mouse major periph
33	40	28.8	45	1	Y12940	Amino acid sequenc
34	39.5	28.4	132	1	P70411	ORF 8 gene product
35	39.5	28.4	244	1	W29149	Human high affinit
36	39.5	28.4	244	1	W75918	Human beta subunit
37	39	28.1	205	1	R05232	Human lymphotoxin
38	39	28.1	1093	1	R56979	Human myotonic dys
39	39	28.1	206	1	R82672	CD45 associating p
40	39	28.1	1989	1	R99640	Peripheral nervous
41	39	28.1	1977	1	R99641	Peripheral nervous
42	39	28.1	467	1	W00382	Bacillus cellulase
43	39	28.1	467	1	W05731	Cellulase. Cellula
44	39	28.1	81	1	W30082	Rat persephin. GDN
45	39	28.1	164	1	W38705	S. pneumoniae phos

ALIGNMENTS

RESULT 1

ID   P50161
AC   P50161;
ID   P50161; standard; Protein; 166 AA.
DT   27-NOV-1991 (first entry)
DE   Sequence encoded by cDNA clone exhibiting multi-lineage cellular
DE   growth factor activity.
KW   Growth factor; multi-lineage; mast cell; haematopoietic.
OS   Mus musculus.
PN   EP-138133-A.
PD   24-APR-1985.
PF   29-SEP-1984; 111677.
PR   04-OCT-1983; US-539050.
PA   (SCHE) SCHERING CORP.
PA   (DNAX-) DNAX RES INST MOLEC.
PI   Yokota T, Lee FD, Rennick DM, Arai KI;
PT   New Poly:peptide(s) having growth factor activities - are prepd.
PT   by recombinant DNA procedures
PS   Claim 5; Fig 1; 64pp; English.
CC   The cDNA is derived from messenger RNA isolated from a mouse T-cell
CC   line after activation with concanavalin A. The cDNA was cloned by
CC   incorporation into a plasmid vector, which then transformed into
CC   E.coli. The plasmid vector also contained DNA segments from the SV40
CC   virus, permitting expression of the cDNA after transfection into a
CC   mammalian host cell, such as monkey COS-7 cells. The polypeptide
CC   includes a potential leader sequence of about 19 AAs.
SQ   Sequence 166 AA;

Query Match   100.0%;   Score 139;   DB 1;   Length 166;
Best Local Similarity   100.0%;   Pred. No. 7.2e-14;
Matches   30;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30
        |||||||||||||||||||||||||||||||
Db   1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

RESULT 2

ID   W10638
AC   W10638;
ID   W10638 standard; Peptide; 30 AA.
DT   23-JUN-1997
DE   Murine IL-3 signal sequence.
KW   LERK; ligand for eph-related kinase; ERK; NLERK2;
KW   receptor protein tyrosine kinase; cell proliferation;
KW   cell differentiation; cell survival; nerve cell; interleukin-3;
KW   IL-3; signal peptide; protein secretion.
OS   Mus sp.
PN   WO9704091-A1.
PD   06-FEB-1997.
PF   19-JUL-1996; AU004263.
PR   20-JUL-1995; AU-004263.
PR   27-NOV-1995; AU-006847.
PR   22-DEC-1995; AU-007299.
PR   05-FEB-1996; AU-007890.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PI   Nicola NA;
WPI; 97-134632/12.
DR   Nucleic acid mol. encoding ligand for eph-related kinase - useful
PT   for treatment of. pref. neuronal, cells to increase survival,
PT   proliferation and differentiation
PS   Example 3; Page 30; 71pp; English.

Sequence Comparison "C" ←

CC A peptide sequence (W10638) comprises the signal sequence of
CC mouse interleukin-3 (IL-3). Mammalian expression vector pEF-BOS
CC was engineered to contain DNA for the IL-3 signal sequence and for
CC a FLAG epitope. PCR fragments (see also T60970-72) coding for
CC mature or soluble NLERK2 (see also W10637), a novel ligand for
CC eph kinase (LERK), were cloned into the vector to allow prodn. of
CC recombinant NLERK2 proteins in transfected COS cells.
SQ Sequence   30 AA;

Qy 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30
      ||||||||||||||||||||||||||||||
Db 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

Query Match   100.0%;   Score 139; DB 1; Length 30;
Best Local Similarity   100.0%;   Pred. No. 1e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
ID W09823
AC W09823;
DT 15-JUL-1997 (first entry)
PN WO9715663-A1.
PD 01-MAY-1997.
PF 23-OCT-1995; AU-006135.
PR 23-OCT-1995; AU-006668.
PR 22-DEC-1995; AU-007276.
PR 09-SEP-1996; AU-002208.
OS Mus sp.
KW Interleukin-3 receptor; cytokine; allergy; asthma; therapy.
DE Interleukin-3; IL-3; signal peptide; NR4; haemoprotein receptor;
DE Interleukin-3; IL-3; signal peptide.
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
PI Hilton DJ, Metcalf D, Nicola NA, Willson T, Zhang JG;
WPI; 97-259018/23.
DR DNA encoding animal haemopoietin receptor which interacts with
PT interleukin-13 - useful to treat asthma, allergy or condition
PT exacerbated by IgE production
PS Example 2; Page 56; 93PP; English.
CC Using PCR, a derivative of novel haemoprotein receptor NR4 cDNA was
CC generated which encoded the interleukin-3 signal peptide (W09823)
CC and an N-terminal FLAG epitope tag (W09824) preceding the mature
CC coding region (Thr27-Pro424) of murine NR4 (see also W09821). The
CC PCR product was cloned into the mammalian expression vector
CC pEF-BOS.
SQ Sequence   30 AA;

Qy 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30
      ||||||||||||||||||||||||||||||
Db 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

Query Match   100.0%;   Score 139; DB 1; Length 30;
Best Local Similarity   100.0%;   Pred. No. 1e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
ID W56260
AC W56260;
DT 16-SEP-1998 (first entry)
DE Construct containing mature interleukin-13 binding protein.
DE Therapeutic; IL-13 mediated condition; allergy; asthma; diagnosis;
KW autoimmune disease; antibody; immunotherapy.
OS Homo sapiens.
PN WO9810638-A1.
PD 19-MAR-1998.
PF 10-SEP-1997; AU0591.
PF 10-SEP-1997; AU-005374.
PR 27-FEB-1997; AU-002262.
PR 10-SEP-1996; AU-002262.

PA (AMRA-) AMRAD OPERATIONS PTY LTD.
PI Hilton DJ, Nicola NA, Simpson RJ, Zhang J;
WPI; 98-207062/18.
N-PSDB; V22701.
DR New isolated interleukin-13 binding protein - used to develop
PT products for therapy e.g. for allergic conditions such as asthma or
PT for diagnosis or detection
PS Example 14; Page 52-53; 69PP; English.
PT The IL-13 binding protein and related therapeutic molecules can be used
CC in the antagonism of at least one IL-13 activity. They can be used for
CC treating IL-13 mediated conditions such as certain allergic conditions
CC such as asthma or to inactivate locally administered IL-13 after IL-13
CC treatment. The products can also be used as diagnostic agents, e.g. for
CC detecting autoimmune diseases. The antibodies can also be used for
CC immunotherapy and may also be used as a diagnostic tool.
SQ Sequence 359 AA;

Qy 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30
      ||||||||||||||||||||||||||||||
Db 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

Query Match   100.0%;   Score 139; DB 1; Length 359;
Best Local Similarity   100.0%;   Pred. No. 1.7e-13;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
ID W55017
AC W55017;
DT 29-SEP-1998 (first entry)
DE Murine IL3 signal sequence.
KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW cell survival; therapeutic; neuronal proliferation; drug screening;
OS Mus sp.
PN WO9811225-A2.
PD 19-MAR-1998.
PF 11-SEP-1997; G02479.
PF 11-SEP-1997; G02479.
PR 11-SEP-1996; AU-002246.
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
PA (DZIE/) DZIEGLEWSKA H E.
PI Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI Zhang J;
WPI; 98-260970/23.
DR New isolated haemopoietin receptor - used for developing products
PT for modulating proliferation, differentiation and survival of cells,
PT e.g. neuronal cells
PS Claim 29(1); Page 54; 182pp; English.
CC The mouse IL3 signal sequence was used to study the expression of mouse
CC NR6.1. NR6.1 is a form of NR6 a novel Haemopoietin receptor (HR).
CC Interaction between the novel HR and a ligand facilitates proliferation,
CC differentiation and survival of a wide variety of cells. The HR and it's
CC derivatives can be used for modulating the activity of the receptors e.g.
CC to regulate development, maintenance or regeneration in an array of
CC different cells and tissues in vitro and in vivo. They can be present in
CC therapeutics used for modulating neuronal proliferation, differentiation
CC and survival. The products can also be used for detection and diagnosis,
CC e.g. for cancers or predisposition to cancers, or for drug screening.
SQ Sequence   30 AA;

Qy 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30
      ||||||||||||||||||||||||||||||
Db 1 MVLASSTTSIHTMLLLLMLFHLGLQOASIS 30

Query Match   100.0%;   Score 139; DB 1; Length 30;
Best Local Similarity   100.0%;   Pred. No. 1e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

Run on: September 19, 1999, 11:38:06 ; Search time 3012.52 Seconds
(without alignments)
12491.011 Million cell updates/sec

Title: US-09-037-657-38
Perfect score: 11832
Sequence: 1 gcggccgctgcagtgattac...........tatgtgctgggtgtgggggga 11832

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl:*

```
 1: gb_ba1:*
 2: gb_ba2:*
 3: gb_om:*
 4: gb_ov:*
 5: gb_pat:*
 6: gb_ph:*
 7: gb_pl1:*
 8: gb_pl2:*
 9: gb_pri:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_st:*
14: gb_sts:*
15: gb_sy:*
16: gb_un:*
17: gb_vi:*
18: em_fun:*
19: em_htg:*
20: em_hum1:*
21: em_hum2:*
22: em_in:*
23: em_om:*
24: em_or:*
25: em_ov:*
26: em_pat:*
27: em_ph:*
28: em_pl:*
29: em_ro:*
30: em_sts:*
31: em_sy:*
32: em_un:*
33: em_vi:*
34: gb_htg1:*
35: gb_htg2:*
36: gb_in1:*
37: gb_in2:*
38: gb_ba1:*
39: gb_ba2:*
40: em_hum3:*
41: em_hum4:*
42: gb_pr4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	11830.4	100.0		11832	5	A70408	A70408 Sequence 38

ALIGNMENTS

c	No.	Score	Query %	Length	DB	ID	Description
	2	6561.6	55.5	6663	5	A70398	A70398 Sequence 28
	3	737.2	6.2	938	5	A70386	A70386 Sequence 16
	4	737.2	6.2	1930	5	A70392	A70392 Sequence 22
	5	607.4	5.1	40668	11	AC003112	AC003112 Human DNA
	6	280.4	2.4	1629	5	A70382	A70382 Sequence 12
	7	280.4	2.2	1673	5	A70384	A70384 Sequence 14
	8	254.4	2.2	834	5	A70388	A70388 Sequence 18
	9	220.2	1.9	1716	11	AF059293	AF059293 Homo sap1
c	10	190.8	1.6	1391	5	A70394	A70394 Sequence 24
	11	141.2	1.2	560	5	A70393	A70393 Sequence 23
	12	132	1.2	33613	12	AF060868	AF060868 Mus muscu
c	13	127.8	1.1	114898	12	AF139987	AF139987 Mus muscu
c	14	127.6	1.1	157684	12	AC004093	AC004093 Mus muscu
	15	127.6	1.1	128026	12	AC004155	AC004155 Mus muscu
	16	127.6	1.1	213245	12	MMHC425018	MMHC425018 Mus muscu
	17	126.2	1.1	21984	12	MMAJ3128	MMAJ3128 Mus muscu
c	18	126	1.1	12300	12	AJ003128	AJ003128 Mus muscu
c	19	125.8	1.1	170351	12	U84903	U84903 Mus musculu
c	20	125.4	1.1	124230	12	MMBAC394	MMBAC394 Mouse BAC
	21	125.2	1.1	118235	12	AC003060	AC003060 Mus muscu
	22	124.6	1.1	245439	12	AC005259	AC005259 Mus muscu
	23	123.8	1.1	273800	12	MMHC135G15	MMHC135G15 Mus muscu
	24	123.8	1.1	4784	12	AF050157	AF050157 Mus muscu
	25	123.8	1.0	6089	12	AF100956	AF100956 Mus muscu
c	26	123.6	1.0	144723	12	MMA3CA212	MMA3CA212 Mus muscu
	27	123.6	1.0	170965	12	MMCABL3	MMCABL3 , complet
	28	123.6	1.0	57327	12	AC005818	AC005818 Mus muscu
c	29	123	1.0	91638	34	AC005302	AC005302 Mus muscu
	30	123	1.0	120479	35	AC003018	AC003018 Mus muscu
	31	123	1.0	128143	12	AC002489	AC002489 Homo sapl
c	32	122.8	1.0	94459	35	AC006945	AC006945 Mus muscu
c	33	122.8	1.0	128008	35	AC003063	AC003063 Mus muscu
	34	122.6	1.0	140554	12	MMTSXDNA	MMTSXDNA Mus musculus
c	35	122.2	1.0	43576	12	AC000096	AC000096 Mus muscu
c	36	122	1.0	179436	35	AC003399	AC003399 Genomic s
	37	122	1.0	182721	35	AC003994	AC003994 Mouse cos
c	38	122	1.0	136687	12	AC003066	AC003066 Mus muscu
c	39	121.8	1.0	158405	12	AC006507	AC006507 , complet
c	40	121.8	1.0	85931	12	MMHC310M6	MMHC310M6 Mus muscu
c	41	121.4	1.0	151218	35	AC005816	AC005816 Mus muscu
c	42	121.2	1.0	143	5	A70390	A70390 Sequence 20
c	43	121.2	1.0	29117	12	AC007049	AC007049 , complet
c	44	121.2	1.0	4101	12	MMCABL2	MMCABL2 Mus musculu
	45	121	1.0	14318	12	AF136179	AF136179 Mus muscu

RESULT 1

A70408

```
LOCUS       A70408   11832 bp   DNA       PAT   07-MAY-1999
DEFINITION  Sequence 38 from Patent WO9811225.
ACCESSION   A70408
NID         g9477483
VERSION     A70408.1  GI:474683
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 11832)
  AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
  TITLE     A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
  JOURNAL   Patent: WO 9811225-A 19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES             Location/Qualifiers
     source          1..11832
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT     2447 a   3367 c   3298 g   2720 t
ORIGIN
```

Query Match   100.0%;   Score 11830.4;   DB 5;   Length 11832;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 11831;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy      1 gcggccgctgcagtgattacttcaccgctgtggtggcaccccacccgcgggccgctgagtgga      60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGGCCGCTGCAGTGATTACTTCACCGCTGTGGTGGCACCCCACCCGCGGGCCGCTGAGTGGA      60

Qy     61 tttttccgtgtgggggatgtgtgaagaagtttaggggagaactcttctgcaccgatgtggaacta     120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTTTCCGTGTGGGGGATGTGAAGAAGTTTAGGGGAGAACTCTTCTGCACCGATGTGGAACTA     120

Qy    121 ggaatgcaggttcgttccgttcccaaaggaacaacctcttcccataagcccactcat     180
Db    121 GGAAATGCAGGTTCGGTTCCGTTCCCGTTCCCAAAGGACACACCTCTTCCCATAAAGCCCACTCAT     180

Qy    181 aagggctccctgcacgcgctccgggagaacatcccatatcccagcagatatgatagt     240
Db    181 AAGGGCTCCCTGCACGCGCTCCGGGACAATCCCCATATCCCAGCAGATATGATAGT     240

Qy    241 tgagaaggaccagagggccggagacttcctccgggacatcccctgcttctgtttccccacgcccctgc     300
Db    241 TGAGAAGGACCAGAGGGCCGGAGACTTCCTCCGGGACATCCCCTGCTTTCTGTTTCCCCCACGCCCCTGC     300

Qy    301 acgaaacgagactacacgcatgggagaggtggcatgaaggcttaggtgggtgggatcggtag     360
Db    301 ACGAAACGAGACTACACCGATGGGAGAGGTGGCATGAAGGCTTAGGTGGGTGGGATCGGTAG     360

Qy    361 gacccatgcaccagagaaaggaaagactggttggcaacttcaaactctctgtggagaaga     420
Db    361 GACCCATGCACCCAGAGAAAGGAAAGACTGGTTGGCAACTTTCAAACTCTCTGTGGAGAAGA     420

Qy    421 aggggctgaaaagaagggctcaagtgctgtttgtttgtgtggcggaccaaag     480
Db    421 AGGGGCTGAAAAGAAGGGCTCAGGTACTGCTCAAATGTGTGTGGCGGACCAAAG     480

Qy    481 tgggtatggggcccgtaagagagggcgggaaggttgatagaaggaaggatcccggtagact     540
Db    481 TGGGTATGGGGCCCGTAAGAGGGGGCCGGGAAGGTTGATAGGAAGGATCCCGGTAGACT     540

Qy    541 ggagggggatcctggaaaagcacacagagagaaagctggggaggtctgggagctggttaaggt     600
Db    541 GGAGGGGGATCCTGGAAAAGCACCAGGGCTAGGAACCCATTCGGAGTTAAGGGT     600

Qy    601 acaggatcccacagtgaagggggcccagagccttgggaggcggaccagagggaggtcc     660
Db    601 ACAGGATCCCAGATGAGGGGGTGGGAAGCCTGGGGCCGCGGGGGGCGGAGGTCC     660

Qy    661 cacggctgttgggaaagagtgtggggggcttcgcgcagggaggatggggacgttcaggagtg     720
Db    661 CACGGGCGTGGTGGGGAAGAGTGGGGGCTTCGCGCAGGGGATGGGGGACGTTCAGGAGTG     720

Qy    721 gtaactgggggccacgatcaaccccccccccgggggggcgggggcgtggggaaggccg     780
Db    721 GTAACTGGGGGCCCACGATCAAAACCCCCCGGGGGGGCGGGGGCGTGGGGAAGGCCG     780

Qy    781 gtgcggggggcccacagggggcccggggggggcggggggggggcccggg     840
Db    781 GTGCGGGGGGCCCACGATCAAACCCCCCAGGGGGCGCGGGGGGGCGGGG     840

Qy    841 cgggggagcgggcgcgcattagcgccttgtcaattcggctgctgatcttgctccggccttt     900
Db    841 CGGGGGCGGCAGCCGCGCGCATTAGCGCCTTGTCAATTCGGCGCTGCTTGCTCCGGGGA     900

Qy    901 cgctgtccgcgggccgcgcgtcgaccgccggcacccgcccccaatcgcacccccccagat     960
Db    901 CGCTGTCCGCGGGCCGCGCGTGACCGCCGGCACCCGCCCCCAATCTGCACCCCCCGA     960

Qy    961 cgccccgcgcccaataccgcggttgcagtcacgcgccccccatgcccgc     1020
Db    961 CGCCCCGCGCCCATACCGGCGGTTGCAGTCCACCGCCCCCCATGCCCGC     1020

Qy   1021 gggtcgccggcggccccgtcgcccaatcgcgcgcgggcgcgcgccgcgcggccgcgctgtcctcggct     1080
Db   1021 GGGTCGCCGGCGGCCCCGTCGCCCAAATCGCGCGCGGGCGCGCGCCGCGCGGCCGCGCTGTCCTCGGCT     1080

Qy   1081 gtggtcgcctctgttgctctgtcctcggggttcgggcgcggggcggatgggaagccccgtga     1140
Db   1081 GTGGTCGCCTCTGTTGCTCTGTCCTCGGGGTTCGGGCGCGGGGCGGGATGGGAGGCCCCGTGA     1140

Qy   1141 gtaccggtcgcgccctgtcccccaccctcccaccccaccgggaagcgggatccccgggggggt     1200
Db   1141 GTACCGGTCGCGCCCTGCTCCCCACCCTCCCGTGTCCCTCGGGTGTCGGGGGGGCGGGGGGT     1200

Qy   1201 agtccgcggggatggaagaagggcgagcgcgcccaactggagctcccggggaccaaaggaa     1260
Db   1201 AGTCGCGCGGGGATGGAAGAAGGGGCGAGCGCGCCCCACCTGGAGCTCCCGGGAACAAAGGAA     1260

Qy   1261 gcgggccctgcgggcgccctcaccttgggggcttcatggggaccacccagcctccaag     1320
Db   1261 GCGGGCCCTCGGGGCGCCCTCACCTGTGGGGCTCAGGGACCACCAGCCTCCAAG     1320

Qy   1321 agtaccccgttatacataaagaagccttcttattgtatccccttgtcgaagcgtgtctgcc     1380
Db   1321 AGTACCCCGTTATACATACAGAGGCCCTCTTATCGTATCCCCTTTGCGAGGCTGTCTGCC     1380

Qy   1381 aggctcagtttgaaggaacatcgcagtgtcctgggaccccctcttcagggttgctgggac     1440
Db   1381 AGGCTCAGTTTGAAGGACATCGCCAGTGTCCCAGCCCCCTTCAGGGTTGCTGGGAC     1440

Qy   1441 gctctcgggggcacgcctgtcttggatattcagagcggaaggaagcctcacggcgg     1500
Db   1441 GCTTCGGGGGCACGCCTGTGTCTTGGAATATCAGAGCGGAAGGCCTCCCTGGCCGG     1500

Qy   1501 gggcgcacgcttgggtcgttgggtcgtgtcggcaaagtggggtccccccat     1560
Db   1501 GGGCGCCAGCCTTGGGTCGGGTGGGTTGGGTGCAAAGTGGGGTCCCCCCAT     1560

Qy   1561 gaagtgatgatcccgggggaggggttatcgtgcggcaaagtcccctggc     1620
Db   1561 GAAGTGATGGATCCCCGGGGAGGGGTTATCGTGCGGCAAAGTCCCCTGGC     1620

Qy   1621 atgcgggcccgggtccggacttgcctctcccgtggggttcggggaccctcctgt     1680
Db   1621 GCTTCGGGGCACGCCTGTGTCTTGGATATCAGAGCGGAAGGCCTCCCTGGCCGG     1680

Qy   1681 ctataggcagactccatgctctttggtactctcctgaagtcctccacactgttgaaac     1740
Db   1681 CTATAGCAGACTCCATGCGTTTGGTATCCTCGAAGTCCTCCACTGGTGGCTCACAAC     1740

Qy   1741 cggtctcattcaggctgcgctgggttgaagcctctagcgactgaaattcgtgaggag     1800
Db   1741 CGGTCTCATTCAGGCTCGCTGGGTTGAAGCCTCTAGCGACTGAAATTTCGGTGAGGGA     1800

Qy   1801 cgagagcaaggccgtgtccgggcaggcccgagacttcaattgttctaagggaccccagt     1860
Db   1801 CGAGAGCAAGGCCGTGTCCGGGCAGGCCCGAGACTTCATTGTTCTAAGGGGACCCCAGT     1860

Qy   1861 ggggggtcaagtctccgagaggaggaactcctggcttgagcgcccc     1920
Db   1861 GGGGGTCAGCTGCCGAGAGGAGAATCCCACTGTCCCAGGAGGAACTCGTGCCTTGAGCGCC     1920

Qy   1921 ataccccaacgacaacatcccagcaagatgcggtggttcccacatccccaggga     1980
Db   1921 ATCACCCAACGACACATCCCCGCCAAGGAATGCGGTCTCCCACATCCCCAGGGA     1980

Qy   1981 cacaccaaagacacaacaaagtagcgccccactggcttatgtcccgtcaccctccga     2040
Db   1981 CACACCCAAAGACACACAACAAAAGAGCCCCACTGGCTTATGTCCCGTCACCCTCCGA     2040

Qy   2041 cgcgggctgcagcccagatgcgtattcgcacacatcgcgcgctgcaatccaatcctct     2100
Db   2041 CGCGGCGCTGCAGCCCAGATGCGTATTCGCACACCATCGCGCGATTCCATCCTCT     2100
```

```
Qy  2101 acacacacacacacacacacacacacacacagacacgcacac                              2160
Db  2101 ACACACACACACACACACACACACACACACACACAGACACGCACAC                          2160

Qy  2161 acacgcacgcacacacacgcacccgcacacccgcacctggtcccacacattcacacagggag          2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ACACGCAGGCACACACACGCACCCGCACACCCGCACTCGTGTCCCACATTATTTCACACAGGGGAG     2220

Qy  2221 gcaacaccggggtacgcatatggttgagtgcactggagactcttcagg                        2280
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 GCAACACCGGGGTACGCAATATGGTTGAGTGCACTGGAGAGTCCATTCCTCAGG                  2280

Qy  2281 acagaaacctacaactctcttgcccaagcggcgctggcactctgggc                         2340
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 ACCCCATCCGGAGACACAGGCCACACCGCAGGGGCCTGGGCTGGGC                          2340

Qy  2341 tagtagtctcttgtgtcaagtttgtctgtggagactggagccctcagggggg                    2400
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TAGTAGTCTTGTGTCCAGTTTTGTCCGAGTTTCCGGGGGG                                2400

Qy  2401 acaggaaccttccgggagaacaggcccaaacgctggggctgctgcctct                       2460
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 ACAGGAACCTACACTCCTGCTTGCCCAAGGCTGCTGATGTGGTTGACACCCG                    2460

Qy  2461 gaccttccgggagaagtggttgctgtctgccaagcctgggtagttctttgaatgccacaata         2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GACCTTTCCGGGAGAGTTGGGTGTTGCTGCCCAAGGCCGCTGGGCAGGTGGTGGTGACACCC        2520

Qy  2521 gcgctaagctttgttctgctgcagagacgcagaggcgaaggcgaaagg                        2580
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 GCGCTAAGCTTTGTTCTGTTCGGGGGCTCGAGATCGCAGAGGCAACAGGTGGGGG                2580

Qy  2581 tgggcgctgtgtttttctttttaagggggagaaattaaataagagttctcaacactc             2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 TGGCCCTGTGTTTTTTTCTTTTAAGGGGAGAAATTAAATAAGAGGTTCACACCTC               2640

Qy  2641 tgcaatctgttgtacctgtgtcttaacctgaccagccagccggtggggtcgta                   2700
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TGCAATCTGTTTGTACTTACCGTGTCTTAACCTGACCAGCCGTGGTCTA                       2700

Qy  2701 aaagtgtatgcagtaccagggggacaaggggttgggggttatggtggatg                      2760
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 AAAGTGTATGCAGTTGCCAGGGGTACCAGAGATGTGGGGCGGGTGTGGGATG                    2760

Qy  2761 gaggccaccttccccgttgggccttcaggaatctcacactttccccttttaaaacactg            2820
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 GAGGCCACCTTCCCGTTGGCCTTTCAGGGGAATCACACTTTCCCTTTAAAACACATG             2820

Qy  2821 gtgtctttttaataacggcagcaactccgcattgggaaggggaaataagg                      2880
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 GTGTCTTTTTAATAAACGGCAGCAACTCCGCATTGGGAAAGGGGAAATAAGCTTGTATA           2880

Qy  2881 ggcccggtccttgttggaggaaaaaaggaggggtgtctctcca                             2940
         |||||||||||||||||||||||||||||||||||||||||
Db  2881 GGCCCCGGCCTTTGTGTGGAAAGGAAAGGGGTGTCTCCTCCA                             2940

Qy  2941 agttgcttggccccattgtttgtggaagccaagaggggactctg                            3000
         |||||||||||||||||||||||||||||||||||||||||
Db  2941 AGTGGCTTTGGGCCCAATTGTTTGTGGAAGCCAAGAGGGGACTCCTCTGG                      3000

Qy  3001 tacttccagagtcaagcttctccagccccagagaaaacgtcttccctgtttattatggag          3060
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TACTTCCAGAGTCAGGCTTCTCAGGCGTCCGAGCCCAGAGAAAACGTCTTCCCTGTTTATGGAGG     3060

Qy  3061 gaatacaggggaaggggggtgccaagtgggactaccgttctctgaggggactgtaccagtgctcg     3120
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 GAATCAGGGAAGAGGGGGTGCCAGGTGGCACTACGTTCTGCTGAGGGACTGTACCAGTGCTCG       3120

Qy  3121 aaggagaaagcttggggcttgcccccctccccaagccacgaagggcgcagctgctaggc            3180
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 GAATCAGGGAGAAGGGGGTGCCAGGTGGCACTACGTTCTGCTGAGGACTGTACCAGTGCTCG        3180

Qy  3181 aaggagaaagcttgggggcttgcccccctccccaagccacgaagggcgcagctgctaggc          3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 AAGGAGAAAGCTTGGGGCTTGCCCCCCTCCCCAAGCCACGAAGGGCGCAGCTGCTAGGC          3240

Db  3241 TAGTGTGGGTAAAAAGGGGCGATTACTCCCCCCCCCAGCCAGGAGGAGTCCCCTGGC             3300

Qy  3301 cagacaaatgctcggggagggaggaggggtgtgatcattgccaagaggaagttcaagtt           3360
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CAGACAAATGCTCGGGGAGGGAGGAGGGGTGTGATCATTGCCAAGAGGAGTGCGAGACAGTGG      3360

Qy  3361 ggtccagggtgggggacagtgcctcccccctcctgttgagggggtggcgcagaggaacaagcc       3420
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 GGTCCGGGGTCGGGCAGTGCCTCCCACCTGCTGAGGGCCTGAGGAGGAACGGGT               3420

Qy  3421 gggtgggcgggggttagaacgctggacgttcccagcatctgcgggggaattctgaattc          3480
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GGGTGGGCGGGGGTTAGAGAGCCCTGGGGACGTCCACGTTCACTGCGGGGAATTCTGAATTA       3480

Qy  3481 gcgggcgggctcggggaccttcggggcctcccgggggggccccgcgcgcctcgtctg            3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 GCGGGCGGGCTCGGCTGCCTCCTGGACCTCCCGCGGCCCCGCGCGCCTCGGCTG              3540

Qy  3541 gcctgctcctccggggactgtccctccctggctctcggagacctcggggacctgggacacccc     3600
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GCCTGCTCCTCCTCGGGGACTAGAACCATGTTGGGCCCTGAGGCCCTGGGACAAGC            3600

Qy  3601 cccaaatgcaactgccgatttgcagagggccgaagaccgcctccccaaggccaaatttg           3660
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 CCCAAATTGCAACTGCGATTCGGAATTGGCATTCGCAAGACCCCTCTCCGGAAGGCCAAAATTTG    3660

Qy  3661 gcagcctctgttctccgagcctctttggaaatcggtttgtttttgttttttcc              3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GCAGCCTCCGTGTCTCCGAGCCTCTTTGGAAATCGGTTTGTGTTTTTGTTTTTTCC            3720

Qy  3721 gggcatggaggagacgtccagggggcgtgcagttccgaagcccctcccccaaatttg           3780
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GGGCATGGAGGAGACGTCCAGGGGCGTGCAGTTTAGCTTCTTAATAGGAAACCTGGGGTGGGGT     3780

Qy  3781 gcagcctctgttctccgagcctcttggaaatcggtttgtttttgttttttcc               3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GCAGCCTCCGTGTCTCCGGACGCTCTTTGGAAATCGGTTTGTGTTTTTGTTTTTTCC           3840

Qy  3841 aatactcttttccctccatcccagactgtttcctccaagggttaactgtcc               3900
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 AATAACTCTTTTCCCTCATCCCAGACCCTTACTGCAGTTTCCTCCCCTTGGTAACTGTCC        3900

Qy  3901 tgcagtctcccctaacctttgctttcttttctccttaccaccaagggttcacatggagtccca     3960
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 TGCAGTCTCCCTAACCTTTTCTTTCTTACCCCACGGGCCTTGCACATGGAGTCCCA            3960

Qy  3961 cctctccccttgcccaactgggtctccagccttactggaatttggcttctttgagcagggtgcc   4020
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 CCTCTCCCCTTGCCCAACTGGGTCTCCAGCCTTACTGCAATTGCCCTTGGTAACTGTCC         4020

Qy  4021 caggcgcctcctgacacacaggggttgttgtgtatgagcccacctctcccctccccccttt      4080
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CAGGCGCCTCCTGACACACAGGGTTTACTTTCCTCTCTGGTTAACTGTCC                 4080

Qy  4081 ctcttttgctcttgagacttaattctttttttttggcttttttgagacagggtt            4140
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 CTCTTTTGCTTCTTCTGAGACTTAATTTTTTTCTTTTTCTTTTTGGCTTTTTGAGACAGGGTT     4140

Qy  4141 tctcctgtacagcccctgggctgcctggactcattctgtagaccaggctagcctcaaactc    4200
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TCTCTGTACAGCCCTGGGCTGCCTGGACTCATTCTGTAGACCAGGCTAGCCTCAAACTC       4200

Qy  4201 acaaacctaccctgcctctgcctttcagtgtctggcactaaagatgtgggccaccacaact    4260
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 ACAAACCTACCCTGCCTCTGCCTTTCAGTGTGGCACTAAAGATGTGGGCCACCACAACT       4260

Qy  4261 agtagttaagtgtttgctgtcttattcctatagtgacctcagtcctggcatattg          4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   4261 AGTAGTTAAGTTGTTTTGCTGTGTCTTTATTCCTATAGTGACTTCCTGGCATATTG 4320

Db   4321 taggggatcggatggatggaatggatggatggatggatggatggaacaag 4380
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAGGCGGATGGATGGATGGAAAGGGATGGATGGATGGTTGGATGGAGCAAG 4380

Db   4381 cttgaatcgtcctcaagtgaaaaaagagaacctcagagaactgaattccaq 4440
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 CTTGAAATCGTCCTGAGTGAAAAAAGAGAACTCAGAAGGAGTTAGGTTCCCAG 4440

Qy   4441 ggcagccctgccctcgctcggtgtctcatcaggagctccctgtgaaactcccccacacctcccac 4500
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 GGCACCCTGCCCTGCCTCGCTCGGGGCTCCCTGTGAAAACTCCCCACACCTCCCAC 4500

Qy   4501 caccctgccatcctctgtgtgggctgacaaggccaatggggacacagactc 4560
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 CACCCTGCCATCCTGTGTGCTGACAAGGCCAATGGCCACTGTGGGGGACACTC 4560

Qy   4561 agggaagttgggaatatgttccccctcctcctcaatatcctaggcctaggggc 4620
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 AGGGAAGCTTGGGAATATGTTCCCCCTCCTCCTCAATATCCTAGGCCTAGGGGGC 4620

Qy   4621 ccaggcctagtaggcagcttgggctgccccttaagtgttgggttaggcaagaagggtg 4680
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 CCAGGCCTAGTAGGCAGTTGGGCTGCCCCTTAAGTTGGGTTAGGCAAGAAGGGGGTG 4680

Qy   4681 gtccctcaggtgggtcacggggaaatgaggattgaagctcagcttcacacaggttcac 4740
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 GTCCCTCAGGTGGGTCACGGGGAAAATGAGGCTCAGCCCCTTCCCACAGGTTCAC 4740

Qy   4741 aggaaatgattgtggaagtcaagaaactcctgttgggaacctgatgtgg 4800
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 AGGAAATGATTGTGGGAAGTCAGAGAACTCCCTGTTGGGAGTTGTAGGGCCTTGCACAGTGTGG 4800

Qy   4801 gcttctgttggtgtccttctctgtgtcttgcacagtcccctgtgttgctggga 4860
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GCTTCTGTGTGGCTGTGTCCCTTCTGTGTGGTCCTTCGGTGTGTGTCTGGGA 4860

Qy   4861 tgtgaggagggcacgggggaaatgaggctcagcctcagtcagtgaaagac 4920
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4862 TGTGAGGAGGGCACGGGGTGGGTCACGGGGAAAATGAAGGCTCAGCCTGCTTCCCTTCACGGTTCAC 4920

Qy   4921 ccaacaaggctcaacctctcctcactcttgatgttggcctcaca 4980
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 CCAACAGGGCTCACCTCTCCTCGTCCTCTGGGACCAGGCTCTCACTGTATGCACAGATTGGCTCACA 4980

Qy   4981 tttgattccctttccttggtctcctgggatgacaaaaacatttaccagg 5040
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 TTTGATTCCCTTTCCTTCTTTGGTCTCCTGGGATGACAAAACATTTACACAGGTGATTTTACA 5040

Qy   5041 ttttagatatgttccattctccagaaacacacacttgtgaggttagtatcagtgaaaggac 5100
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTTTAGATATGTTCCATTCTTCTCCAGAAAACACACTTGTGAGGTTAGGTTAGTATCAGTGAAAGGAC 5100

Qy   5101 accccaggagacaagacaaaatttgggaaaaatgggtaagccaccagtgcttga 5160
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 ACCACCAGGACAGACACAAGAATTGGAGAAGGAAGGAAATTGGCCATGCCTTGA 5160

Qy   5161 tggcttatagtgtaatcccagaaacctctgggacgctcagcgcaggcaggaagaagttccaaga 5220
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 TGGCTTATGTGTAATCCCAGAAACCTCTGGGACGCTCAGGGCAGTTCCAAGA 5220

Qy   5221 cagtgtgttctaggtaatgagcacctgtcaagaaaagaaagataagacaagaaa 5280
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 CAGTGTGTTCTAGGTAATGAGCACCCTGTCAAGAAAAGAAATAAAGACAAGAAA 5280

Qy   5281 atgtttatatAGGGCTGTTCAGACAGCTTGGTCGTGGGTAAGGGGACAGTTGCCTCCAATCAAGATGA 5340
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 ATGTTTATAGGGCTGTTCAGACAGCTTGGTGCGTGGGTAAGGGGCACTTGCCTCCAATCAAGATGA 5340

Qy   5341 cctcagccccatccctaggaatccattgggtagaagggaaaagcaaaactccagctgctgacc 5400
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 CCTCAGCCCCATCCCTAGGAATCCATTGGGTAGAAGGAGGGAAAGCAAAACTCCAGCTGCTGACC 5400
```

```
Qy   5401 tccatacatgttgctccaaatgtgcacacacaggggagacataatcaattaataggatgta 5460
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 TCCATACAATGTTGCTCCAAATGTGCACACACAGGGAGACATAATCAATTAATAGGATGTA 5460

Qy   5461 tttgcttagatttgagtaggcatttatgactgatgttttaaaatttatttgatttat 5520
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 TTTGCTTAGGATTTGAGTAGGCCATTTATGACTGATGTTTAAAAATTTTTATTTTAT 5520

Qy   5521 gaaaatattacctgtttgtatttgtttgtttgtttgtttgtttgtttgagacagg 5580
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 GAAAATATACCTGTTTGTATTTGGTTTTGGTTTTTGAGACAGG 5580

Qy   5581 gcttctctgtgtagtcctggctgtcctggaactcactctgtagaccaggctggcctga 5640
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTCTCTGTGTAGCCCTGGCTGTCCTGGAACTCACTCTGTAGACCAGGCTGGCCTTGA 5640

Qy   5641 acctcagagaaatccgacctgcttctgtatgtgctaggattattaaaggtgtgcactgcat 5700
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 ACTCAGAGAAATCCGCCTGCTTCTGCCTCCCCAAGTGCTGGGATTAAAAGGTGTGCACTGCCAT 5700

Qy   5701 tcagcaaaaattgcatacttaaacccccagtattgttggaggaaggcagaagaactaaggttga 5760
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 TCAGCAAAAATTGCATACTTTTAACCCCAGTATTTGGGAGGCAGACTAAATGTGTGA 5760

Qy   5761 attccaggctagcacaaggatacacagagtgagaccccattttcactgtctagc 5820
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 ATTCCAGGCTAGCACAAGGATACACAGAGTGAGACCCATTCTTCTTCTTTCACCTCCCCAAAAAC 5820

Qy   5821 acctcagaaatccgacctgctctgtgtagctcagtccaagtgtgggcacgtaaagtgtccaagg 5880
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 ACTCAGAAAATGTATTTTGTAGCTTGTGTTGTATGTACAGTGTGTGTCACGTAAAGTGTCCAAGG 5880

Qy   5881 acaacttgtgaagttctctctcccgttcacacttctaagtcccgaaataagggtctc 5940
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 ACAACTTGTAGGAAGTTCTCTCTCCGTTCACAGTCTCTGAAATTGATACTTCAG 5940

Qy   5941 aggcttagcacaagtctcttctcttatgtgtactgacccattccactggccctggatttgactgat 6000
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 AGGCTTAGCACAAGTCTTCTTTTATGTACTGACCATTTCACTGGGCCCTGGATTGACTGAT 6000

Qy   6001 ccaaaatgtatttgtcttgtatgtacagtgtgttgcttagctagctcagcacgtaaaatgtccaag 6060
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCAAAATGTATTTTGTAGCTTGTGTTGTATGTACAGTGTGTGCTTAGCTAGCTCAGCACGTAAAATGTACATATGCC 6060

Qy   6061 aaggtcatctgagtgctgtcggtactcttgtctccaccccaaagtgttggaatgatctcag 6120
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 AAGGTCATTCGTGAGTGCTGCTGGTACTCTTGCTTCCACCCCAAAGTGGTGGAATGATACTCAG 6120

Qy   6121 gcagcacttctctgggagggggctggcccttgatcttttgttgcctcagcttcaat 6180
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GCAGCACTTCTCTGGGAGGGGGCTGGCCCTTGATTTTTGTTGCCTCAGCTTCAAT 6180

Qy   6181 gagtgcttgggtctctgtgttctttcttttctctgtgaaatggttgaaacctgttcaa 6240
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 GAAATTAATTTTGAGATAAGGATAAGGTCTCTGTTCTTTATCTCTGTGAAAATGGTTGAAAACTATTGAAACTCCC 6240

Qy   6241 gacttcctgactctgctattacatcaggcaacgtgggactgagggactgggctcatccat 6300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 GACTTCCTGACTCTGCTATTACATGGACACCCCAGGACACCCCAT 6300

Qy   6301 gcctaccaagtgtcgtcttgacccccagaccaagcctgtaatcagccatccccc 6360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 GCCTAACAAAGTGTGTCGTCTTGACCCCCAGACACCAGGGACCCCAC 6360

Qy   6361 ccttctcatcggctcctccaagctactgtgtcctccaaatggacaccctgggggc 6420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 CCTTCTCATCGGCTCCTCCCTGCAAGCTACCTGCTCTATACATGAGACACACCTGGGGGC 6420

Qy   6421 caccgcctgagggggctctactgggacctcaatgtcgcccgcctcgagctgtcccg 6480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 CACCGCTGAGGGGCTCTACTGCGACCTTCAATGGTCGCCGCCTCTGAGCTGTCCCG 6480
```

```
Qy  6481 cctccttaaacacctcaccctggccctggctaacccttaatgggtccaggcagca 6540
         ||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  6481 CCTCCTTAAACACCTCCACCCTGCCCTGGCTAACCTTAATGGGTCCAGGCAGCA 6540

Qy  6541 gtcaggagagacaatctggttgtgtcacgcccgacacggcagccagccctcctgcct 6600
         |||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db  6541 GTCAGGAGAGACAAATCTGGTTGTGTCACGCCCGAGACGGCGAGCAATTCGGCTCCTGCCT 6600

Qy  6601 ctatgttggctgtaagtgggcgtgggggggttggcaatgaca 6660
         |||||||||||||||||||||||||||||||||||||||||
Db  6601 CTATGTTGGCTGTAAGTGGGCGTGGGGGGGTTGGCAATGACA 6660

Qy  6661 gattagaggctggtgtcttctgtcttgggcagagccagccttcttcttcaggcagca 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 GATTTAGAGGCCTGGGTCTGGGTGTCACGGAGCCAATCGGCTCCTGCCT 6720

Qy  6721 catggtcataccacagcacagagacacttcagagacagatagatagtgggttggcaatgaca 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 CATGGTCATACCCACGCACAGCCATTGCAACTCTTAGGACACAGCTGTGGCTGCAATGACA 6780

Qy  6781 tgtgttaccccacagctcttagaaagctgtcatgtttcctgtagtgtcccccctgagaagc 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 TGTGTACCCCACAGTTTAGAAAAGTGTCAGTTTTCCTGTGAGTGTCCCCCTGAGAAGC 6840

Qy  6841 ccttaaacatcagttcgctgctgtgttcccggaacatgaagtgccgctggacaccgg 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 CCTTTAACATCAGCTGCTGTGTTCCCGGAACATGAAGGAATCTCACGTGCGGTGGG 6900

Qy  6901 gtgcacacgggagagacaattcttacatacccaaagtacaacgtgcccc 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 GTGCACACGGGGAGAGACAATTCTTACATACCACAAGTACAAGGTGGG 6960

Qy  6961 taccagccaagccttgctctgtgtgtgtgaacttctggcatacttacctctctgatcaaatatg 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 TACCCAGCCAAGCCTTGCTGTGTGTGAAGAACTTACCTCTCTGATCAAATATG 7020

Qy  7021 ttcctgtttattgaactcacacgtctgtgtacggtacggtcaggataa 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 TTCCTGTTTATTGAACTCACAAAAGGGACTCTCGCACCTTCGCACATTAA 7080

Qy  7081 cacatgtgaggagtaccacacagttcttacatacccaaactccatcaaagtacaacgtgagctggc 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 CACATGTGAGGAGGTACCACACTGTGGGCCCTCACTCATGCCATAATCCCAAGGACCTGGC 7140

Qy  7141 cccttcactccttatgatctcggtggttgaagccaccaatcgcctaggggctcaagcaagatc 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 CCTCTTCACTCCCTTATGAGATCTGGGTGGTGAAGCCAACCAATCGCCTAGGTCAGCAAGATC 7200

Qy  7201 tgatgtcctccacactggatgtctccggacgtgggttgagccccagtggtcccacctgtgttct 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TGAAGTCCTCACCACTGGANTCCTGGGTCTGGACTGGAGCCCCAGTGTCCACCTGTGTTCT 7260

Qy  7261 gccctagacctttataggcgcgcccccccaatcccccacagacttttgttcctctagag 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 GCCCTAGACCTTATAGGGCGGGCCTCCCCATCCCCACAACTTTTTGGTTCTTCTCTAGAG 7320

Qy  7321 gtcttagccacagccacggtggttgcaggacagtggttgttcatcaataacttaatgcaaagac 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 GTCTTAGCCACAGCCACGGTGTTGTTCATCAATAACTTAATGCAAAGAC 7380

Qy  7381 tttcccccaagacagtcaagaatttctccctcaccacacacacacacaca 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 TTTCCCCAAGACAGTCAAGAATTTCCCCTCACCACACACACATACACACA 7440

Qy  7441 ctctgcagagacaacctgggcctgacacacctccctcctacacgccaggaagg 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 CTCTGCAGAGAACACCTGGCCCTCCTCTTCTCAGCCCAGCTGTTCAGGGAGG 7500

Qy  7501 gagtcctcaggggactgagagagagcgcccccaggttctgagaggcc 7560
         |||||||||||||||||||||||||||||||||||||||||
Db  7501 GAGTCCTCAGGGGACTGAGAGGGAGGGCCCCCAGGTTCTGGAGGGAAGCCGAGGCC 7560

Qy  7561 ttgagctggggggggcgaggggttggaggcacgaggtgatcctcctgagacaact 7620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  7561 TTGAGCTGGGGGGGGGGGGGGGGCGGAGGGGGTTGGAGGCACGGGAACTGAGACAACT 7620

Qy  7621 gggggcctaatcttaattagggtgttcccagcccaaagcagcagcctgggcattaacccttcaa 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGGGTTCCTGGCTAAACTTAATTAGGGTGGTTCCCAGCCCAAAGCAGCCTGGGCATTAACCCTTCAA 7680

Qy  7681 gtgcctcactgaagactcaggggagggagagactcaagcttgttactctcttcctcatggtccccagga 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 GTGCCTCACTGAAGACTCAGGGGAGAGACTCAGGGCATTAACCCTTCAA 7740

Qy  7741 gggttcctgggtgtgccccctggctcattcccacacccagagggtttctgttcctctggcatc 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GGGTTCCTGGGTGTGCCCCTGCCTCATTCCCACAAGAGGGTTTGTGTTCCTCTGGCATC 7800

Qy  7801 taaccctcagttgtgctctggtcagctggccttctcaccacccagaggttttgtcttcctggcatc 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TAACCCTCAGTTGTGCTCTGGTCAGCTGGCCTCATTCCCACAAGAGGTTTGTGTTGGTTAATGTA 7860

Qy  7861 caaggcatcagaggtggacatggggatcacatagggagtgggcaaatagcactc 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CAAGGCATCAGAGGTGGACATGGGGATACAGGTGGGCAAATGCACTC 7920

Qy  7921 aaggtggggtggtgatatacaataaagcttgtcacctgaccgcctgacgccaaagcctactctcatgat 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AAGGTGGGGTGGTGATATACAATAAAGCTTGTCACCTGACCGCCTCAGAAAGCCTACTCATGAT 7980

Qy  7981 gatcacaattgttgacttcactctgggacatgtagtcagacccctgagtcaaaacacagac 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 GATCACAATTGTTGACATCACTCTGGGACAGTGTAGTGAGACCCCTGAGTCAAAACACAGAC 8040

Qy  8041 agtagcttaagagtcagcttgtgcttaactactggacctaatacttcaggggcctaatagttgctgg 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 AGTAGCTTTAAGAGTCAGCTTGTGCTTAACTACTGGAACTCAGGGGCCTAATAGGTGCTGG 8100

Qy  8101 gtgaatgtcgcctcacctccgtttagtgagatctcgcgctaatctcactccacccagctgg 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 GTGATGCTCGCGCCTCACCTCACTCCGTTTAGTGAGACTCACGGGGCCTAATAGGTCATGG 8160

Qy  8161 gatcacaattgttgacattcaactctgggaggatgtgtctcccattgagggacccgg 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 GTGGGGCTGCTCTGTCCCCTTGAGGGCAGGAANTGTGTCTCCCATCAGGATAGGACCCG 8220

Qy  8221 tggtagagcaactgtgctgctggctgttcttctggaattaaatgacagtaactctatcaggc 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 TGGTAGGAGCAACTGTGCTGGCTGTTTCTGGAATATTAAATGACAGTAACTATCAGGC 8280

Qy  8281 ctgggtgagtagctaacagagggtgtggggggcgttggctggaaaacagagatagggtcattagg 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGGTGGTAGGCAGCAACTGTGTGCTGTTTCTGGAANATTAAATGACAGTATATCAGGC 8340

Qy  8341 agccactgacctagcctgagaattacacacaactgggggggggtctggaaaacgagcatagggtcatagg 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGCCACTGCGCCTAGCATTACACACCACTGGGTTCTCACAGCGCATTCTCACCTCACCCAAGC 8400

Qy  8401 agtcctcagaactgggagcactgtgccagccaattaatgccagcattaatgccaagc 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 AGTCCTCAGAACTGGGAGCACTGTGCCAGCCAATTTAATGCCAGCAATTTAATGCCAGCATT 8460

Qy  8461 aggggggaggcagaggcagaaggagtctcctcgagttcaagaccagccatcctgaatttacacataaag 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 AGGGGAGGCAGAGGCAGAGGATCTCTGAGTTCAAGGCCATCCTGAATTACATAAAG 8520

Qy  8521 agctccaggccaggggtgcgcagccaagtaaaacctgttctcaaaaaaacacttag 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 AGCTCCAGGCCAGGCCGGTGCGCAGTAAAAAACCTGTTCTCAAAAACAAAGAACCATGTTTAG 8580

Qy  8581 tgacaggcttgctccaccaccaccaccccgacccacgtgcaacgtgagcc 8640
         ||||| |||||||||||||||||||||||||||||||||||||||||||
Db  8581 TGACAGGCTTGCTCCACCACCCCCCAGGTGGCGACCGCCCACCACGTGCTGCACGTGAGCC 8640

Qy  8641 gcgttggggcctgggagggaccagctgagtgtcgctgggtcctcaccacccagctctcaagg 8700
         |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  8641 GCGTTGGGGCGGGCCTTGGAGGACCCAGTGAGTGTGGCCTGGGGTCTCACCCAGCTCTCAAGG 8700

Qy  8701 atttcctcttccaagccaagtaccagatccgcgtaccgcgtggaggacagcgtggactgga 8760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 ATTTCCTCTTCCAAGCCAAGTACCAGATCCGCTACCGCGTGGAGGACAGCGTGGACTGGA 8760

Qy  8761 aggtgccccgtcccgtcccgccccgaccccgcaatctgactcctcct 8820
         |||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AGGTGCCCCGTCCCCGCCCCGGACCCCGACCATTGACTCCTCCCT 8820

Qy  8821 caccgtgcaggtgggtgacgtcagcaacaccagaccttcctcggg 8880
         |||||||||||||||||||||||||||||||||||||||||||||
Db  8821 CACCGTGCAGGTGGGTGGGATGACGTCAGCAACACCAGACTCCTTCGGG 8880

Qy  8881 gcccggcaccgttacttcgtcaagaagtcgcgttgtaacccattcggggatctaattcagtgg 8940
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 GCCCGGCACCGTTTACTTCGTCCAAGTCGCGTTGTAACCCATTCGGGGATCTAATTCAGTGG 8940

Qy  8941 aaaggcgggaatcctgaaggag 9000
         |||||||||||||||||||||
Db  8941 AAAGGCGGGAATCTGGAGCGAGTGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGG 9000

Qy  9001 tgagcacctcctcaaggggctcggctggcccatggaatcccaccaatccatcctgttccttccccc 9060
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGAGCACCTCTCAGGGGTTGGCCTGGCCCATGGGATCCCCAATCCATCCTGTTCTTTCCCC 9060

Qy  9061 ccaccccttctttttgagacagcgtcttcaggtagcgcatgggccttaaattcagttatg 9120
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 CCCACCCTTTTTTTGAGACAGCGTCTTCAGGTAGCGCATGGGCCTTAAATTCAGTGAATG 9120

Qy  9121 tagtcaaggatgacctcagtcctggtctttttgtctccctcacttagagacaatggccagt 9180
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9121 TAGTCAAGGATGACCTCAGTCCTGGTCTTTTTGTCTCCCACTTAGAGACAAATGGCCAGT 9180

Qy  9181 ggccatcaccactcttgggagaacttgtagtctattagcctgtgtcatttggtgtgaca 9240
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9181 GGCCATCACCACTCTTTGGGAGAACTTGTAGTCTATTTAGCCTGTGTCATTTGGTGTGACA 9240

Qy  9241 gatggagtacacagtgtgacctccttgtaagagaacagcgctgagaactcagttatgtctt 9300
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9241 GATGGAGTACAACAGTGTGACCCTTGTAAGAGAACTGAGGCTGAGTGTTTTTATGGCTT 9300

Qy  9301 aatatcctagagctcctctagaggttaacttatttatataaaaataagagacaatggccagt 9360
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9301 AATATCCTAGGCGTCTCTACTCCTGGCCACCCCTCCAAACCCTTAAAAGGTAACCTAGGCAC 9360

Qy  9361 tcacatggtcccacagaacctttgtcacacaacctagccacagccatgcctgcctac 9420
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9361 TCACATGGTCCCACAGAACCTTTTGTCACACAACCTATATGACACTGCCCTGTGCCCTAC 9420

Qy  9421 cacctaaggtctcactgctgtgcccaaccctccaacccttaaaagtaacctaggcagc 9480
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9421 CACTAAGGTCTCTTACTCCTGGCCACCCCCCAAACCCTTAAAAGGTAACCTAGGCAC 9480

Qy  9481 cttaaatatttgcaatccccatcagcctcttgaatgctcaggacgcattaacc 9540
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9481 CTTAAATATTTGCAATCCTCCTACCTCTTGAATGCTCAGGAGAACAGGGCATTAACC 9540

Qy  9541 caagtttcctctctctggtccctttctctttaagctgtggggaggggcctaaagtgactt 9600
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9541 CAAGTTTCTCTCTTCTCTGGGTCTCCTTTAAGCTGTGGGGAGGGGCCTCTCC 9600

Qy  9601 aatgtctgaagacttccgagcccatggaatctctctaatatgaaatattattgcataa 9660
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9601 GTCCTGAACACTCTCCGAGACCCATGTCCTTTTCTAAAGTGGCTAATATGAAATATTTGCATAA 9660

Qy  9661 aatgtctggcctcagtttcccaagttgtttaggcagcaacagtgtccaagacac 9720
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9661 AAtGTCTGGCCTCAGTTTCCCGAGACTCTCGAGGTTTAGGCAGCAACAGTGTCCAAGACAC 9720

Qy  9721 ttcattattgcaggcagtatataagagaagaagctcccaatcccccaaccccgcttcctccggtcc 9780
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  9721 TTCATTATTTGCAGGCAGTGTATAAGAGAGAAGAAGCTCCCAT 9780
```

```
Qy  9781 ctaagacagcaatacttctacacttctactgaaactgaaactcctcgcagacgcgaatatgctcacttaa 9840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9781 CTAAGACAGAATACTTCTACACTGAAACTCTCGCAGACCGGCATATGCTCACTTTAA 9840

Qy  9841 tgatgatgaaatatgggagaactgaggtcccgaggagattccttggaggaagaagggtcaaa 9900
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  9841 TGATGATGAAATAATGGGGAAACTGAGGCTCCGAGAGATTCCTCGGAGGAAGAGGGTCAAA 9900

Qy  9901 accagcgtccaggagagctccgagggctcccacccccgggcccctccccaagtctttctggttgcgg 9960
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  9901 ACCAGCTCCAGGGAAGCTCTCCAGCCCCATCCGTCGCCGAGTATTCCTCGGGCCTGGGCGG 9960

Qy  9961 gagtgaaccaacagctgggaggggaggggaggggagctgggaggagctttggctgtctgcccagc 10020
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  9961 GAGTGAAACACAGCTGGGAGGGGCTGGAGCCTGGGAGCTTTGGGCCTCCGGGC 10020

Qy 10021 acctgcgattcttgcacgggacggcaggggggtccgtccgccgagagacagctgaaggaag 10080
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10021 TGCTGGGGAGGGGGATGGCGGCCGCCAAGGGGCTGGGCTTCCCACTCAGTCCTC 10080

Qy 10081 ccggggggtaggggtctggggggggcaggggtctgttgtccccagggaagcttgttgccaggg 10140
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10081 CCGGGGGTAGGGGTTGGAGGGCAGCCAGCAGCCCCGAGAGGACTGAAGGAG 10140

Qy 10141 cctgtcaggcggagtccccagtgttttatggcttgaagcgcgacgatgtccttatccggctggcc 10200
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10141 CCTGTCAGGCGGAGTCCCCAGTTTTATTTATGGCGTGGGGAGCTTTATCCGGCC 10200

Qy 10201 ttaatataactacggtttttaaaaattgctttttgttataatggtgtgtgccccagtta 10260
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10201 TTAATATAACTACGGTTTTTAAAAATTGCTTTTGTAAATGGTGTGTGCCCTGAGCGGC 10260

Qy 10261 cagcccacccaatccaacagtcaccccgtgaaaggttagagaaaacttgttgagtaggtctcttcccaatgt 10320
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10261 CAGCCACTCCATGTCAACACACGTGAAGGTTGTGAGAACTTTGTTGAGTAGGCTCCTTG 10320

Qy 10321 ttctgtgtctgtctctattctgtcattccacctttccccagagccctttttttttatgctt 10380
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10321 TTCTGTGCGTGTCTGTCTCTCTRATTCACTTTGCTCCAGAGCCCTTTTTTAGCTT 10380

Qy 10381 gggactaggggctgcggcgacaagagcaattactgagtcaatctcgccaagccccctca 10440
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10381 TTAATATAACTACGGTTTTAAAAATGCTTTTGTAAATGGTGTGCTGCCCTCAGTGC 10440

Qy 10441 gtgcacacacacacgtgaaggttgttagagaaacttgttgagtaggtctcttcccaatgt 10500
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10441 GTGCCAACACACACGTGAAGGTTTGTTGAGAACTTTGTTGAGTAGGCTCCTTCCCACATGT 10500

Qy 10501 gggactaggggctgcggcgacaagagcaattactgagtcaatctcgccaagccccctca 10560
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10501 GGGGACTAGGGGCTGCGGCGACAAGAGCAATTACTGAGTCAATCTCGCCAAGCCCCCTCA 10560

Qy 10561 cttcccatcgttttggatagttcatgaggtaatcgaggttgccagctgtctaattttcg 10620
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10561 CTTCCCATCCTGTTTTGGATAGGTRATCGAGGTTTGCCAGCTGTCTTAATTTCG 10620

Qy 10621 tagtatctgcctcagcctaccaagtgctgtgtaccacgttgtgtgggagggctctcc 10680
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10621 TAGCTATTCCTGCCTCAGCCTACCAGTGCTGTGTACCACGTTTGTGGGAGGGCTCTCC 10680

Qy 10681 tccccagtgtctgtgggggtacacacgtcccaagaatcctgctcttctgctttctctgtc 10740
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10681 TCCCCAGTCGTCGTGGGGGTACACAGTCCCAAGAATCCTGCTCTTCTGTGTCTTTAGTT 10740

Qy 10741 tgccccttgctttgtccggtggtcctcggaagtctcggagtctccatgagttc 10800
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 10741 TGCCCCTTGCTTTGTCCGGTGGTCCTCGGAAGTCTCGGAGTCTCCATGAGTTTC 10800

Qy 10801 ctttctgaccgaataactcggtttttacctccccactgatttgactcccctccttgttctt 10860
         ||||||||||||||||||||||||||||||||||||||||||||||||
Pb 10801 CTTTCTGACCGAATACTCGGTTTTACCTCCCACTCGATTTGACTCCCCTCTTTGCTTGTCT 10860
```

RESULT   2

LOCUS   A70398   6663 bp   DNA   PAT   07-MAY-1999
DEFINITION   Sequence 28 from Patent WO9811225.
ACCESSION   A70398
NID   g4774676
VERSION   A70398.1   GI:4774676
KEYWORDS   .
SOURCE   unidentified.
  ORGANISM   unidentified
             unclassified.
REFERENCE   1 (bases 1 to 6663)
  AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
             Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
             Kikuchi,Y.
  TITLE   A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
  JOURNAL   Patent: WO 9811225-A 19-MAR-1998;
             NICOLA NICOS ANTONY (AU)
FEATURES   Location/Qualifiers
     source   1..6663
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT   1462 a   1852 c   1715 g   1634 t
ORIGIN

Query Match   55.5%; Score 6561.6; DB 5; Length 6663;
Best Local Similarity 99.6%;
Matches 6628; Conservative 0; Mismatches 19; Indels 5; Gaps 5;

Qy   5186 tggacggctgaggcgaggaggaatt-ccaagttccaagacactgtg-ttctaggtaatgagac 5243
           |||||||||||||||||||||||||| ||||||||||||||||||| |||||||||||||||||
Db   12 TGGACGGCTGAGGCGAGGAGGAATTCCCAAGTTCCAAGACACTGTGTTTCTAGGTAATGAGAC   71

Qy   5244 cctgtcaagaaaagaaaataaagagacacaagaaaatgtttatagagctgttgagacagc 5303
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 CCTGTCAAGAAAAGAAAATAAAGAGACACAAGAAAATGTTTATAGAGCTGTGAGACAGC   131

Qy   5304 ttggtgggtaaggtgggcacttggctcccaccacttcagtgcccatccctcaggaatc 5363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 TTGGTGGGTAAGGTGGGCACTTGGCTCCCAACAAAGATGACTCAGCGCCATCCCTAGGAATC   191

Qy   5364 catggtagaaggtgagagggaagccaaactc-cagctgtgacctccatacactgtgccaatgtg 5422
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   192 CATGGTAGAAGGAGAGAAAGCAAACTCGCACCTGTGACCTCCATACACTGTGCTCCAAATGTG   251

Qy   5423 cacaacacaagggagacatataatcaaattaatactaggatgtatttgttagttttgaggca 5482
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 CACAACACAAGGGAGACATAATCAAATTAATACTAGGATGTATTTGTTAGTTTTGAGTAGGCA   311

Qy   5483 tttatgactgatgttttaaaattttattttatttgaaaatataacctgtttgtatt 5542
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 TTTATGACTGATGTTTTAATTTTAATTTTTATTTGATTTATGAAAATATAACCTGTGTTTGTATTT   371

Qy   5543 ggtttggttttggtttggagtgttttgttttattttgagacaggggctctctcgtgtggct 5602
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 GGTTTGGTTTTGGTTTGGAGTGTTTTGACTTTGGTTTTATTTGAGACAGGGTCTCGTGTAGCCCTGGCT   431

Qy   5603 gtccttggaactccactctgtagacacaggctgcctcgaaatccgcctgcttgt 5662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 GTCCTTGGAACTCACTCTGTAGACCAGGCTGGCCTTGAACTCAGAAATCCGCCTGCTTGT   491

Qy   5663 gctccccaagtcgcttagattaaaggtctgcactgccattcagcaaaattgcatacttaa 5722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   492 GCTTCCCAAGTGCTTAGATTAAAGGTGTGCACTGCCATTCAGCAAAATTGCATACTTTAA   551

Qy   5723 ccccagtcatttgggaggcagaggcagactaatgtggaattccaaggctagccaaggatac 5782
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 CCCCAGTATTTGGGAGGCAGAGGCAGACTAATGTGGAATTCCAGGGCTAGCCAAGGATAC   611

Qy   5783 agagtgagacacctattcctaccctccccccccaaaaatgtatttttgtgcttgt 5842
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   612 AGAGTGAGACACCTATTCTTACCCTCCCCCCCCCAAAAATGTATTTTTGTGCTTGT   671

Qy   5843 gtatgtacatgtgtgttgcagcacgtaaatgtccaaggacaacttgtagaagttctctcc 5902
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GTATGTACATGTGTGTTGCACCACGTAAATGTCCAAGGACAACTTGTAGAAGTTCTCTCC   731

Qy 10861 ccatcgcctggcattgcattgcattctctctgtgggtgactctgtccacaccttgacaccttcc 10920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10861 CCATCGCCTGGCATTGCATTGCATTCCCTCTGGGTGACTCTGTCCACACTGACACCTTTCC 10920

Qy 10921 caactttccccagccagccgaagctggtcctgtatggagcgcgccgtcccgcggcctcct 10980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10921 CAACTTTCCCCAGCCAGCCGAAGCTGGTCGTCCTGGGTAATGGAGGCCGCGCCCGTCCCGCGGCCTCCT 10980

Qy 10981 gctggcgccgcccaacactgccgctcatcttcttagtgagcgccgccccggcgccgggcg 11040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10981 GCTGGCGCCGCCCAACACTGCCGCTCCTCCATTCTCTTTAGAGCGCCGCCCCGGCGCCGGGCG 11040

Qy 11041 gggtgtgtccagcgcgcgcgggcgggcgagcccagcctcgggcgccgagctcaagc 11100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11041 GGGTGTGCCAGCCGCGGGGGCGCCAGCCCAGCCTCGGGCGCCGGGCGGGCGGGCGAGCTCAAGC. 11100

Qy 11101 agttcctcggctggctcaagaagcaacgatactgctcccggccccgggccggagctcacg 11160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11101 AGTTCCTCGGCTGGCTCAAGAAGCAACGCACGGCAATACTGCTCCGAACCTTAGTTTCCGCCCTGTACG 11160

Qy 11161 accaagtggcgttgcttggatgcacaagagacaccgaaaaccagtggaggaaagttgg 11220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11161 ACCAGGTGGCGCTTGGATGCACAAGAGTCACACAAGAAACCAGTTAGGGAAAGTTGG 11220

Qy 11221 gggaggcgttgcctgcggtggggtaaagagacagaggaggaagcagcctc 11280
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11221 GGGAGGCGTTGCCTGCGGTGGGGGTAAAGGAGGCGGGTGAGCGACTCGAGGAGCGAGCCTC 11280

Qy 11281 cacaacaccgcactctctctccaagaacagcagagagaggaccaggggtgagcgacctc 11340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11281 CACAACACCGCACTTCTTCTTCCAAGCACAGCAGAGAGAGGACCCGGGGTGAGCGACCTC 11340

Qy 11341 gggtgcgcgagaggtaaggtgggtctgggtgagtgggccttacaagcagtctagaataagc 11400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11341 GGGTGCGCGGAGAGGTGAAGGGTGGGTCTGGGTGAGTGGGCCCTTACAGCAGTCTAGGTGATGAGGC 11400

Qy 11401 cctttccctccttcggtgttgctcaaagggatcctttagtgctcattcaccactgca 11460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11401 CCTTTCCCTCCTTCGGTGTTGCTCAAAGGGAATCTCTTAGTGCTCAATTCAACCCACTGCA 11460

Qy 11461 aagagccccaggttttactgacttgctgaaggtgccagtgcttaatgtggcct 11520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11461 AAGAGCCCCAGGTTTTACTGCATCATCAAGTTGCTGAAGGGTCCAGGGTTCAGGGTTGCCT 11520

Qy 11521 ctttcctgcctcagtgcctcccggctaaactctaaggatagggccaatcctcctgctgggt 11580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11521 CTTTTCTGCCTCCTCGGCTAAACTCTAAGGATAGGCCCATCCTCCTGCTGGGT 11580

Qy 11581 cagacctggaggctcaactggagccccctcgtaccatctgggccaacaagaaaacc 11640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11581 CAGACCTGGAGGCTCAACTGGAGCCCCTCGTACCCATCTGGGCAACAAAGAAAACC 11640

Qy 11641 taccagaggctgggaacaaagagagctcccacaaccacagagtcac 11700
           ||||||||||||||||||||||||||||||||||||||||||||||
Db 11641 TACCAGAGGCTGGGCACAATGAGCTCCCACAACCACAGCTTGGCTCCCACATAGTGGTCAC 11700

Qy 11701 acttggatataccccagtgtggttcaggggcctccaagagtctct 11760
           |||||||||||||||||||||||||||||||||||||||||||||
Db 11701 ACTTCGGATATACCCCAGTGTGGGTGTTATTTGGGGTTCAGGGGCCTCCCAAGAGTCTCT 11760

Qy 11761 ttaaataaataaaggagttcaggtcccgatggccagtgtgttttgggcctatgtgct 11820
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11761 TTAAATAAATAAAGGAGTTGTTCAGGTCCCGATGGCCAGTGTGTTTTGGGCCTATGTGCT 11820

Qy 11821 gggtggggggga 11832
           ||||||||||||
Db 11821 GGGGTGGGGGGA 11832

```
Qy   5903  gttcacagtctaagtcctgaattcaaactaaggtctcaggctacacagtctcctt   5962
           ||||||||||| |||||||||||||||||||||||||||||||||
Db    732  GTTCACAGTCTTAAGTCCTGAATTCAAACTAAGGTCTCAGGCTTCTTCTT        791

Qy   5963  atgtactgaggccattcactggccctgga ttgactgattaatttttgagataagt   6022
           |||||||||||||||||||||||||||||
Db    792  ATGTACTGAGCCCATTTCACTGGCCCTGGATTGACTGATTGAATTAATTTTGAGATAAGGT  851

Qy   6023  ctcttgtaagctctagctcaaacatatgaaactaccccaaggtcatctgagctggt   6082
           |||||||||||  ||||||||  ||||||||||||||||||||||
Db    852  CTCTTGTAGCTCTAGGCTCAAACTATGAAACTATCAAACTCCCAAGGTCATCTCCTATGAGT 911

Qy   6083  acttctgcttccaccccaagtggtcggaatcgtatactcaaggcactcttcctgggg   6142
           ||||||||||||||||||||||||||||||||||||||||||
Db    912  GCTGGCCTTGGCCTTGGCCTAGGCTAGGCTCAAACTATGAAACTCCCAAGGTCATCTGAGCTGGGGAAGGG 971

Qy   6143  gctggccttggcctcagtctcaatgagtgctctggttctcgttgttt             6202
           |||||||||||||  ||||||||||||||||||||||||||||||
Db    972  CCAGGCAGGGTGAGGGGACTTGAAAGTGGGGCTCATCCCATGCTCAACAAAGTGTCGCTTTG 1031

Qy   6203  ctttctttatctgtgaaacacttcaatgagtcatctgaaacat                 6262
           ||||||||||||||||||||||||||||||||||||||||||
Db   1032  CTTTTCTTTATCTGTGAAATGGGTGAACACCTGTTCAAGACTTCCTGACTCTTGAAACAT    1091

Qy   6263  ccaggcaggtgagggacttgaagtgggcttcatcccatgcctaacaaagtgtcgtcttg  6322
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1092  CCAGGCAGGGTGAGGGGCTCATCCCATGCCTAACAAAGTGTCGCTTTG              1151

Qy   6323  acccaggacacagtgtaatcagncccaggacccaccacttctcatcggnctcctcctg   6382
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1152  ACCCCAGACAGTCTAATCAGNCCCAGGACCCACCACTTCTCANTCGGNCTCCTCCCTG    1211

Qy   6383  caaggtaacctgctctatacatggaggacacacctgggggctcccactggt          6442
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1212  CAAGCTACCNCTCTATACAAATGGAGGACACACCTGGGGGCTCAGGACTCTACTGG       1271

Qy   6443  accttcaaatgtcgccgccctgcccctgagcgttgtccogccctcctcaccncctg     6502
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1272  ACCTTCAAATGTCGCCGCCCTGCCCTCCTTAACACCTCCACCCTG                 1331

Qy   6503  gccctggccctgctaaccttaatgggtccaggncagaagnncaatctgggtgtgt      6562
           |||||||||||||||||||||||||||||||||||||||||
Db   1332  GCCCTGGCCCTGGCTAACCTTAATGGGTCCAGGCAGCAGCAACAATCTGGGTGTGT      1391

Qy   6563  cacgccccgagcggcagcaattctgctgtccctcatgtgtctaagtggggc          6622
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1392  CACGCCCGAGCGGCAGCAATTCTGGGTGGCCTCCTGCCTANGTGTTTGGCTGTPAAGTGGGGC 1451

Qy   6623  cccagacaactcagagagagatagatagggggttggcaatgacagatttagagc      6682
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1452  CCCAGACACTCAGAATAGANTAGGGGGGGTTGGCAATGACAGATTTAGAGCTGGGTGTTCTG 1511

Qy   6683  tcctggggcagagccatgggctctcactgcatcgagtcatgtcataccagacaagc     6742
           |||||||||||||||||||||||||||||||||||||||||||
Db   1512  TCCTGGGGCAGAGCCATGGGCTCTCACTTGCATGCAGGTCAATACCCAGACAGGC      1571

Qy   6743  attgcaacttaggagcaacagctgtgggtcactgtccctgtaccocccaagcttagaa   6802
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1572  ATTGCAACTTCTAGGGACAGCTGTGGGTCACTGTCCCTGTGTACCCCACAGCTTTAGAA    1631

Qy   6803  aagcctgtcatgtttcctgttagtgcccctgagaagcccctttaactctgctgggtc   6862
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1632  AAGCCTGTCATGTTTCCTTGTTAGTGCCCTTTAACTCAGCTGTCGGTC              1691

Qy   6863  cggaacaatgaaggatctcacgtcgcgcctggacacggtcgcacagacaagacattctt  6922
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1692  CCGGAACAATGAAGGANTCACGTTCAGTGCCGCTCAGGAAGACCACTCAGTTCTT       1751

Qy   6923  acatacaacaactacctccctcaagtacaagccaagccttgctgtg               6982
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1752  ACATACCAACTACTCCCTCAAGTACAAGCTGAGGTTGGTACCCAGCCAAGCCTTGCTGTG  1811

Qy   6983  tgacttctggcaatactaacttcttctctgatcaaaatatgttcctgttcatcaaaag  7042
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1812  TGACTTCTGGGCAATACTTACCTTCTCTGATCAAAATATGTTCCTGTTTATTGAACTCAAAAG 1871

Qy   7043  ggacttctcgcacctccacaggtggtgcacgtcaggatanaacacactgaaggag      7102
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1872  GGACTTCTGCGCACCTCCACAGGTGGTCASGATAACACATGTGAAGAACACACT         1931

Qy   7103  gtgggccctcaactcatcgcanaactcccccaaggaacctggcccctcttcaactccc  7162
           |||||||||||||||||||||||||||||||||||||||||
Db   1932  GTGGGGCCCCTCACTCATGCCAATAATCCCAAGGACCTTGGCCCCTCTTCACTCCTATGGAGAT 1991

Qy   7163  tgggtgggaagcacacaaatcgcctaggctcagacaagatctgatgttcctacaagg  7222
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1992  TGGGTGGGAAGCACTCACTGCCATACCCAAGGACCTTGGCCCCTCTTCACTCCTATGAGATC 2051

Qy   7223  ctggaagtgggtgagccccacccccaagtgtccacctgtgttctgccctgatatgggc  7282
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2052  CTGGACGTGGGTGAGCCCCANCGCCCTAGGCTGCAAGAACCTTGATATGAGTGTCC      2111

Qy   7283  tccccccaatcccccacccagtcaccctttgttctctagaggtgttgcccctacacgg  7342
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2112  TCCCCCCCATCCCCCACCCAGTCACCCTTGTTTCTCTTAGAGGTCTTCGCCCTCTAGAGGT 2171

Qy   7343  ttgtcaggacatgtggttcttgttcataactaacttgcaaagaacttcaccaccatgg  7402
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2172  TTGCAGGACAGTGGTTGTTCATCCATAACTTAATTGCAAAGAACTTTTTCGGTTCTTCTA  2231

Qy   7402  ttttccctcccacacccccaacaacacacacaaccttctcagagacccggcc        7461
           |||||||||||||||||||||||||||||||||||||||||
Db   2232  TTTTCCCTCTTCCCACCACCCAACAACACACACAACCTTCTCAGAGAACCNGGCC        2291

Qy   7462  tgacccacctcccttctactagccaggtgttcagaaggngagttcagttcctactag  7521
           |||||||||||||||||||||||||||||||||||||||||
Db   2292  TGACCACCTCCCTTCTACTAGCCAGGTGTTCAGAAGGNGAGTTCAGTTCCTACTAGG    2351

Qy   7522  aggcgcccagtctaagtgccccccagggaagggcccctgagctgggggggggggcga  7581
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2352  AGGCGCCCAGTGTCTGAAGGGCGCCCGAGGTTCAGCGCCGAGACTCAGG             2411

Qy   7582  gggtttggaaggcacgaactggtgatgatcctgagcacacaaactgggtagggtg     7641
           ||||||||||||||||||||||||||||||||||||||||
Db   2412  GGGTTGGGAGGCACGAACTGGATGAANTCTGGGCCATCAAATCTAATTAAGGGTG       2471

Qy   7642  ttcccagccaaaggcagccctgggcatcttaacctcactgagaacttaag          7701
           |||||||||||||||||||||||||||||||||||||||||
Db   2472  TTCCCAGCCAAGGCAGCCCTGGGCATTTAACCCTCTACTGAAGAACTCAGG           2531

Qy   7702  ggagagaatcagcttgtactctcatgttcccccaggaggggggttcctgggtgccctggc 7761
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2532  GGAGAGATCAGCTTGTACTCCTCTCCAANGGTCCCCCAGGAGGGTGGGGTCCCGGC       2591

Qy   7762  tcattcccacaccgagggttttttgttctcctggcaatctaaccccagcctgtgctgt  7821
           ||||||||||||||||||||||||||||||||||||||||||
Db   2592  TCAATTCCCACANCCAGAGGTTTTGNTCTTCCTGGAATCTAAACCCCAGTTGTTGCTCTGT 2651

Qy   7822  ggcttggacagtgccccgtggaggtcctcttgg taatgtacaagaggggtggacat    7881
           ||||||||||||||||||||||||||||||||||
Db   2652  GGCTGGGACAGTGCCCCGTGGCGGGTCTTGGTGTACTACAGAGGGGGACAT           2711

Qy   7882  gggatgggggtacaataggggtagagccaaatatagacecctcaagtgtgggttgataatacaaata 7941
           ||||||||||||||||||||||||||||||||||||||||
Db   2712  GGGATGGGGATACANTACGGGGATGGAGCCAAAATAGCACCTGGATGTGATAATACAANTA 2771

Qy   7942  aagcttgtcacccctgacgctcagaaagcctactcatgagtgatcacac            8001
           |||||||||||||||||||||||||||||||||||||||||
Db   2772  AAGCTGGTCACCCTGACGCTCAGAAAGCCTACTCACATCAATGGTGACTAAC         2831

Qy   8002  tctgggacatgtagtgagaccctagtccaaaacaaacagagtaggttaagagtcaagctt 8061
           |||||||||||||||||||||||||||||||||||||||||
Db   2832  TCTGGGACATGTTAGTGAGACCCTAGTTCAAAAACACAGCACAGTAGCTTTAAGAGTCAGCTT 2891

Qy   8062  gtgacttaatactggaactcaggngctggtgatctcgcctcactccct            8121
```

```
Db  2892  GTGACTAATAACTGGAACTCAGGGCGGCTAATAAGGTGCTGGGGTGGTCTGGCCTCACTTCCCT  2951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  8122  gttagttagatctctggcgctaatctcaccacccagctgtgggtgggtgtctgtcccctg  8181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2952  gttagtgagatctcgcgtaattccaccacccagctgggtggggggtctgtccccttg  3011

Qy  8182  aggggcaggaatgtgtgtcttccatcaagagataggaccccgtggtagcagcaactgctg  8241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3012  aggggcagagaatgtgtgtcttccatcacagagataggaccccgtggtagcagcaactgctg  3071

Qy  8242  gctgctttctggaatattaaatgacagtaatccttcatcaggtttgggagttaagctaacagg  8301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3072  gcgtgtttctggaatattaaagacagtaatctatcagggcctgggtgagttagctaacagg  3131

Qy  8302  gtgggggttggttcttggaaaacccagagatagggcacactgggtggtggtagcaactggc  8361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3132  gtggggcctgggtctggaaaaccgagatagggggtcataggacctagagttac  3191

Qy  8362  accactgggtgtcttgtcactaggaaccagtcctcagaactgggagcac  8421
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3192  accactgggttgtcttgtcactaggccattctcaccaagagtcctcagaacttgggagcac  3251

Qy  8422  tgttgccagtcatttaatgccagcattagggaggcagaggaagg  8481
          ||||||||||||||||||||||||||||||||||||||||||||
Db  3252  tgttgccagtcattaatccagcaattttaaatgccagcattagggaggcagaggcagagaagg  3311

Qy  8482  atctctcttgagttcaaaggccacacctgaatttacataaagagctccaggccaggggtg  8541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3312  atctctctgagttcggctgggtctcaggccatcctgaattacataaagagctccaggccaggggtg  3371

Qy  8542  cgcagtaaacctttgtctcaaaaacaaagcatcatcctttgtgaccagggctttgtctccacccc  8601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3372  cggagtaaaacccttgtctcaaaaacaaagacatcttagtgaccaggctttgctcacaccc  3431

Qy  8602  cagtgaccacggaccccgaccgacgtgcacgtgagcgccgttgggggcctgaggacc  8661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3432  cagtgaccacgggaccaccccgaccgcagacgtgcagtgagcgccgttgggggccttggggaacc  3491

Qy  8662  agctgaagtgtgcgctcggggtctcaagagattccttcctctccaagccaagt  8721
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3492  agctgagtgtgcggctcagggtctcaccaccagcttctcaaggattttccttccttccaagccaagt  3551

Qy  8722  accagatccgctaccccgtggagggaagtgtcccgcccccgga  8781
          |||||||||||||||||||||||||||||||||||||||||||
Db  3552  cgtcagcaaccagacctcctcccggtctcgcccgcacgtttacttccgt  3611

Qy  8782  ccggccccctgacccgcacgtgagcgccgttgggggcctgaggatga  8841
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  3612  ccccgccccctgaccccgcatctgactcctcctcaccgttcaggtggttggacga  3671

Qy  8842  cgtcagcaaccagacctcctcgcggctctgcggggccgttcacttcgt  8901
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3672  cgtcagcaaccagacctcctcccggtctcgcccgcacgtttacttccgt  3731

Qy  8902  ccaagtgcgttgtaacccattcgggatcatgggtcgaaaaaaggccggga  8961
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3732  ccaagtggcttgtaacccatcatggggatctatggggtcgaaatctggacga  3791

Qy  8962  gtgggagccacccacccgccccaccctctttttttggag  9021
          |||||||||||||||||||||||||||||||||||||||
Db  3792  gtggagccgcttgtaaaccatcggggatctatgggggtgtgagggctgg  3851

Qy  9022  ctggccatggaatccccaatcccatcctgtcctcaccccacctttttggagg  9081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3852  ctggccatggaatcccaatcccatcgttcttcccccacccttttttgagacag  3911

Qy  9082  cgtcttcaaggtaggacgatcggcctttaaattcaagtgaactccgagc  9141
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3912  cgtcttcaggtagcgcctttaaattcagtgatgtgagtcaagacctcgagc  3971

Qy  9142  tcctggtcttttgtctccacacttagagacaatggccagtggccatcaccacctttgggag  9201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3972  tcctggtctttttgtctccactttagagacaatgggccagtggccatcaccacctttgggag  5052


Db  3972  TCCTGGTCTTTTTGTCTCCACTTTAGAGACAATGGGCCAGTGGCCATCACCACCTTTGGGAG  4031
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  9202  actagccatggagtctgattcatttggttagctgctccacaacagtggac  9261
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4032  actagccatggagtctgattcatttggttagctgctccacaacagtggac  4091

Qy  9262  ctctttggtaaggaactgaagcaaggggctgttcttgtacccaatatcctaggctcttagag  9321
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4092  ctctttggtaaggaactgaagcaaggggctgttcttgtacccaatatcctaggctcttagag  4151

Qy  9322  gttaactttatataaaataagaagactattacagccagttctcacagagaact  9381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4152  gttaactttatataaaataagaagactattacagccagttctcacagagaacct  4211

Qy  9382  tttgtcacacaaccttatagaccacagtgctgctccacacaagagttctctgct  9441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4212  tttgtcacacaaccttatagaccacagtgctgctccacacaagagttctctgct  4271

Qy  9442  ggccaccctcccaccccttaaaggtaagctaggcagcttaatatttgcaatcctct  9501
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4272  ggccaccctcccaccccttaaaggtaagctaggcagcttaatatttgcaatcctct  4331

Qy  9502  acctcagctccttgaatgtccaagaacccaagcttcttctctgggtc  9561
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  4332  acctcagctccttgaatgtccaagaacccaagcttcttctctgggtc  4391

Qy  9562  cccttccttaaggtggggaggcctcaaagagaaagagactccgagcc  9621
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  4392  cccttccttaaggtggggaggcctcaaagagaaagagactccgagcc  4451

Qy  9622  catggaatctgcactcctaatatgaaatatattgcaataaaatgtcttccc  9681
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4452  catggaatctgcactcctaatatgaaatatattgcaataaaatgtcttccc  4511

Qy  9682  cacctgtcaggtttaggcagcacagtcggtccagacacacttcattattttctctgcagtat  9741
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4512  cacctgtcaggttaggcagcacagtcggtccagacacacttcattattttctctgcagtat  4571

Qy  9742  aagaagaagctccatcccacccccgacgctccctcccggtccctcaaagaatacttctaca  9801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4572  aagaagaagctccatcccacccccgacgctccctcccggtccctcaaagaatacttctaca  4631

Qy  9802  ctgaaactgaactctccgcagacgcatagctcacttaatgatgaaataatgggaa  9861
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4632  ctgaaactgaactctccgcagacgcatagctcacttaatgatgaaataatgggaa  4691

Qy  9862  actgaggtccggagagattcctggaggaaggagggtcaaaacaagctccaggaagctctcc  9921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4692  actgaggtccggagagattcctggaggaaggagggtcaaaacaagctccaggaagctctcc  4751

Qy  9922  agccccaatcccaggggcctctccccagggttctgggcttgggtgggaggg  9981
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4752  agccccaatcccaggggcctctccccagggttctgggcttgggtgggaggg  4811

Qy  9982  gctgggagccctgggagcttttggccctgtctcgttcgattcttcctgcaggga  10041
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4812  gctgggagccctgggagcttttggccctgtctcgttcgattcttcctgcaggga  4871

Qy  10042 gccagcaggggctgcgctccgggccgaagagaaccgggggttaggggttggaggga  10101
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4872  gccagcaggggctgcgctccgggccgaagagaaccgggggttaggggttggaggga  4931

Qy  10102 ggtaagcaggggctgtggggcccgaagcttgtgccaggtcagcagagtcccagtt  10161
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4932  ggtaagcaggggctgtggggcccgaagcttgtgccaggtcagcagagtcccagtt  4991

Qy  10162 ttatttaatgggtgaggccgatgtcttatccgctggcctggggggatgggctgct  10221
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4992  ttattttatggggtgtgaggccgatgtcttatccgctggcctggggggatgggctgct  5051

Qy  10222 ggggattggaccaagggctggcttccacactcagtcctcactccatgtcaacc  10281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5052  ggggattggaccaagggctggcttccacactcagtcctcactccatgtcacacc  5111
```

Qy 10282 cgtgcattctctgaggcttatcttgggaacccgcctttgtctgtctgctctctat 10341
         |||||||||||||||||||||||||||||||||||||||||||| | | ||| |||
Db 5112 CGTGCAATTCTCTGAGGCTTATCTTGGGAACCCGCCTTGTTCTGTCGTCTCTAT 5171

Qy 10342 ttctgtcattcaacttcccaggagccttctttttttatgtgctttttaaa 10401
         |||||||||||||| ||| ||||||| ||||||||| || |||||| ||||
Db 5172 TTTCTGTCATTCAACTTTCCCAGAGCCTTTTTTTATGCTTTTTAATATAACTACGTTTTAAA 5231

Qy 10402 aattgctttttgtaataatgtgttgcttctgagcgtgtgcacacaacacacgtgaa 10461
         ||||||||||||| |||| |||| ||||| |||||||||||||||||||||||||||
Db 5232 AATTGCTTTTTGTATAATGTGTTGCCTTCTGTGCGTGTGCACACAACACACGTGAA 5291

Qy 10462 ggttagagaacttgttgttgagtaggtctccttccaacagtgggactaggggctgcgcacaag 10521
         ||||||||||||||||| || ||||||||||||| ||||||||||||||| ||||| |||||||
Db 5292 GGTTAGAGAACTTGTTGACTGAGTAGGCTCCTTCCCAACAGTGGGACTAGGGGCTGCGCACAAG 5351

Qy 10522 agcaattactgagtcatccgccgccccctccacctccaccatgtgggacacaa 10581
         ||||||||||||||||| |||||||||||||||||||||| |||| ||||||||
Db 5352 AGCAATTACTGAGTCATCCGCCGCCCCCTCCACCTCCACCATGTGGGACACAAG 5411

Qy 10582 tcataggttaatcgaaggtaaatcgtgtggctttaattttcgtagtctatctgcctcagccta 10641
         || ||||||||| ||||||||||| |||||||||||||||||||||||| |||||||||||||
Db 5412 TCATAGGTTAATCGAAGGTAAATCGTGTGGCTTTAATTTCGTAGCTATCTGCCTCAGCCTA 5471

Qy 10642 ccaagtgctgtgctaccacacgtgtttgtgggagggctcctccctccccaagtgtctgggggt-aca 10700
         ||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||| |||
Db 5472 CCAAGTGCTGTGCTACCACAGTTTGTTGTGGGAGGGGCTCCTCCCTCCCCAAGTGTCTGGGGGTGACA 5531

Qy 10701 cagtcccaagaatcctctgccctagtccacttagtctccaagccccttgtctgtg 10760
         |||||||||||| |||||||||| |||||||| || || |||||||| |||| |||
Db 5532 CAGTCCCAAGAATCTCTGCTACCAGTTTGTTGCGGAGGGGCTCTCCCCAGTGTCTGGGGTGACA 5591

Qy 10761 tccctagagtcctccgcgcccactttagtctccaagctctgacggaatactcgg 10820
         |||| |||||||||||| |||| |||| |||||||| ||| ||||| |||||||||
Db 5592 TCCCTAGAGTCTCCGCGCCCACTTTATCCATTCCATTCCTTTTCCTTTACCGAATACTCGG 5651

Qy 10821 ttttacctacctccgattgactccctcctcctttgtctccatcgccgttgcattgccа 10880
         |||||||||||||||||||| ||||||| || ||||||||||||| ||||| |||||
Db 5652 TTTTACCTACCTCCGATTTGACTCCCTCCTCCTTTGCTTGTCTCCATCGCCGTTGGCATTGCCA 5711

Qy 10881 ttcctctggttgactctgggtccacactgacaacctttcccagccgaagc 10940
         |||||||||||||||||||| ||| ||||||||||||||||||||||||||
Db 5712 TTCCTTCTGGGTGACTCTGGGTCCACACCTGACAACCTTTCCCAGCCGAAGC 5771

Qy 10941 tggtctgttgtattggaggccgccgtccccggcgcccgtgccccacaact 11000
         ||||||||||||||||||||||||||||||||| |||||||||| | |||
Db 5772 TGGTCTGGTTGTATGGGAGGCCGCCGTCCCCGCGCCCGTGCCCCAAACACT 5831

Qy 11001 gccggctccattcttcttagagagcgccggcccggccccgggccgtgcgccg 11060
         ||||||||||||||||||| ||| ||||||||||||||||| |||||| |||
Db 5832 GCCGCTCCATTCTCTTAGAGGCCGCCCCCGTCCCGCGCCTCCTGCGCGCGGGGGC 5891

Qy 11061 ggccgaagtcacacaagacccgaaccagttaggagaagttggggagggt 11120
         |||||||||| |||| |||||| |||| ||| |||||| |||| ||| |||
Db 5892 GGGCGAGCCCAGCTCGGGCCCCGGGAGCTGGGGGCGGGGGAGGGGGTTGGGATG 5951

Qy 11121 aagcacgcatactgctcgaacctctgatttccgcctgtagcacagtggcgctgg 11180
         |||||||| ||||| ||| ||||| || ||||||| ||||| |||||| ||||||
Db 5952 AAGCACGCATACTGCTCCGAACCTTAGTTCCGCTTGGACCAGTGGGGGGGGGGGT 6011

Qy 11181 caggagtcacacaagaccccgaaaccaggtaggagaagtt 11240
         ||||||||||||||||||| |||||||| |||||||| |||
Db 6012 CAGAGTCACACAAGACCCGAAAACCAGGTAGGAGAAAGTTGGGGAGGCTTGCGTGGGGT 6071

Qy 11241 aaaggagcaggaaggaagagagagagagaccccgg 11300
         |||||||||||| |||||||||||||||| ||||||
Db 6072 AAAGGAGCAGGAGAGGAGAGAGACCCCGGTGAGCAGACGCACTCTTCTT 6131

Qy 11301 tccaagcacaggacgacgaggagggatcctgccctccggggacacggg 11360
         |||||| | ||| |||| |||| || ||| ||||| ||| | |||||||
Db 6132 TCCAAGCACAGGACGAGGGGATCCTGCCCTTCGGGCAGAGCAGAGGGTGCGGGAGGAGGTAAGG 6191

Qy 11361 gggtctggtgagtggggcctacacgacagtctagatgagtgaggcccttcccctccctcggtgt 11420
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 6192 GGGTCTGGGGTGAGTGGGGCCTACACGGGCCAGTGTCAGCAGGTCTAGATGAGGGGCCCTTCCCCTCCCTCGGTGT 6251

Qy 11421 tgctcaaagggatctcttttgttccatttcaccactgcattcaccccac 11480
         |||||||| |||| || ||||| ||| ||||||||| || ||||||||||
Db 6252 TGCTCAAAGGGATCTCTTAGTCGCTCATTTCACCCACTGCAATTCACCCCAC 6311

Qy 11481 catcatcaagttgctgaaggtcaggcttaatgtggcctcttctttctcct 11540
         ||||||||||||||||||||||| |||||| |||| ||| |||||||||||
Db 6312 CATCATCAAGTTGCTGAAGGTCAAAGGGCTCCAGGGTCCTAAATGTGGCCTCTTTTTCTGCCTCAGGT 6371

Qy 11541 gccggctaaaotctaaggataggccaatcctcccgctgggcagacctgaggcctcaactg 11600
         ||||| ||||| |||||||||| ||||| ||||||||||||||||| |||||||||||| ||
Db 6372 GCCGGCTAAAACTCTAAGGATAGGCCCATCCTCCCGCTGGGCCAGACCTGGAGGCCTCAACTG 6431

Qy 11601 aattggagcccctctgtacactctgggcagacaagagaaacctaccagagggctggggaacaat 11660
         |||||||||||| ||||| |||||| ||||||||| | |||| ||||||||| ||| |||||||
Db 6432 AATTGGAGCCCCTCTGTACCATCTCGGGGCAGACAAGAGAAACCTACCAGAGGGGTCAGACACCAAT 6491

Qy 11661 gagtcccacaacacagctttgtccacatgtggtcacacttggaatacccagtgt 11720
         ||||||||||||| | ||||| ||| |||| || |||| || ||||| |||||
Db 6492 GAGCTCCCACAACACAGCTTTTGGGTCCACATGATGGTCACACTTGGATAACCCCAGTGT 6551

Qy 11721 gggtagggttgggtgttattgcaggggcctccccaagagatctcttttaaataaaggagttg 11780
         |||||||||||||| ||| ||||||||| ||||||| || |||||||||||||||||||||||
Db 6552 GGGTAGGGTTGGGTATTTGCAGGGGCCTCCAAGGAAGTCTCTTTAAATAAATAAAGGAGTTG 6611

Qy 11781 ttcaagtccgatggccagtgtgtttgtttggggggcctatctatgtggtgtctgggggt 11832
         |||||| |||| |||||||||||||||||||||||||||| || |||||| | ||||||||
Db 6612 TTCAGGTCCGGATGGCCAGTGTGTTTGGGGGGGGCCTATGTGTGCTGGGGTGGTGGGGGGA 6663

RESULT 3
A70386

```
LOCUS       A70386       938 bp    DNA             PAT
DEFINITION  Sequence 16 from Patent WO9811225.
ACCESSION   A70386
NID         g4774665
VERSION     A70386.1  GI:4774665
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1 (bases 1 to 938)
  AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
  TITLE     A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
  JOURNAL   Patent: WO 9811225-A 19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES             Location/Qualifiers
     source          1..938
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             <1..468
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAB42574.1"
                     /db_xref="PID:e1433693"
                     /db_xref="PID:g4774666"
                     /db_xref="GI:4774666"
                     /translation="GTVFVQVRCNPFGIYGSKKAGIWSEWSHPTAASTPRSERPGPG
                     GVCEPRGGEPSSGPVRRELKQELGLWLKHAYSNLSRLYDGWRAWMQKSHKTRNQV
                     GKLGEACVGGKGAPEEERDGEQPPQHRILLSKHRTRGSCPRADGVRREVRGSG"
BASE COUNT      243 a    245 c    272 g    178 t
ORIGIN
```

Query Match   6.2%; Score 737.2; DB 5; Length 938;
Best Local Similarity   96.2%; Pred. No. 1.8e-145;
Matches 777; Conservative   0; Mismatches 28; Indels   3; Gaps   2;

RESULT 4.

LOCUS   A70392   1930 bp   DNA   PAT
DEFINITION Sequence 22 from Patent WO9811225.
ACCESSION   A70392
NID   g4774670
VERSION   A70392.1   GI:4774670
KEYWORDS   .
SOURCE   unidentified.
ORGANISM   unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1930)
AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
TITLE   A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME

JOURNAL   Patent: WO 9811225-A 19-19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES   Location/Qualifiers
     source   1..1930
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT   375 a   623 c   561 g   371 t
ORIGIN

Query Match   96.2%;   Score 737.2;   DB 5;   Length 1930;
Best Local Similarity   96.2%;   Pred. No. 1.6e-145;
Matches 777; Conservative 0; Mismatches 28; Indels 3; Gaps 2;

Qy   10983 tggccggccgcccaacaactgcgctccattctctttagagcgccagcgccccgggggcggg 11042
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGAGCCACCCCACCGCTGCCTCCACCCCTCGAAGTGAGCGCCCCGGGCCCCGGGGGGG 138

Qy   11043 gtgtgccagccgcgcgggggcggggagccccagcgctcggggcccggttgcggcgcagctcaagcag 11102
                 ||||| | ||||| ||||||||||||||||||||||||||||||||||||||||
Db   139 GTGTGCGAGGCCGCCGGGGGCCCGGAGCCCAGTGTCGGGGCCGGGTGCGGGGAGCTCAAGCAG 198

Qy   11103 ttcctcggctggctcaagaagcacgcatactgctcgaaccttagttctccgcctgtacgac 11162
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 TTCCTCGGCTGGCTCAAGAAGCACCGATACTGCTCGAACCTTAGTTTCCGGCCTGTACGAC 258

Qy   11163 cagtcggccgtggtcttggatgcagaagtcaacagaccgaaaccagtgtaggaaagttgggg 11222
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 CAGTGGCGGGTGCTTGGATGCAGAAGTCACACAGACCGAAAACCAGTAGGAAAGTTGGGG 318

Qy   11223 gaggcgcttgggtgggggttaaagtggaggagagaggaaggaccccgggtgagcagcctcca 11282
                 ||||| ||| |||||||||||||||||||||||||||||||||||||||||||
Db   319 GAGGCTTGCGTGCGGGGGTAAAGGAGCAGAGAGGAAGAGACAACCCGGGTGAGCAGCCTCCA 378

Qy   11283 caacaccgcactctcttcttccaagacacaggacgagagggaaccctcggggcagacggg 11342
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379 CAACACCGCACTCTTCTTCTTCCAAGCACAGGGAGGACGGGGAATCCTGCCCTCGGGGCAGACGGG 438

Qy   11343 gtgcggcgaggagtaagtggggggtctgggtgaaggtgctcagaggaagtcaga 11402
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   439 GTGCGGCGCGGAGGGTAAAGGGGGTCTGGGTGAGTGGGGCCTACAGCAGTCTAGAGATGAGGCCC 498

Qy   11403 ttcccctcctcggttgttgctcaaagggatctctcattcacccactgcaaa 11462
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   499 TTTCCCCTCCTTCGGTTGTTGCTCAAAGGGATCTCTTAGTGTCCTCACCCACTGCAAA 558

Qy   11463 gagcccaggtttactgcatcatcaagttgctgaagggtccaaggttggcctct 11522
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   559 GAGCCCAGGTTTTACTGCATCATCAAGTTGCTGAAAGGGTCCAGGCTTAATGTGGCCTCT 618

Qy   11523 ttttctgccctcaggttctgtcgcggctaaaactcaaagttgctggtca 11582
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   619 TTTCTGCCCTCAGGTCCTGTCGCGGCTAAACTCTAAGGATAGGCCAATCCCTCCTGCTGGGTCA 678

Qy   11583 gacctgagggcttcaacctgaattggagcccctgtgtacc--atctgggcaacaaagaaacct 11641
                 ||||||| ||||||||||||||||||||||||||| ||||||||||||||||| |||
Db   679 GACCTTGGAGGCTCACCTGAAATTGGAGCCCCTTGTGTACTGGGCAACAAATGGGCAAGAAACCT 738

Qy   11642 acca--gaggctgggcacaatgagctcccacaacacagcttggtccacatgatggtca 11699
                 ||||||| |||||||||||||||||||||||||||| ||||||| |||||||||||
Db   739 ACCATGAGGCTGGGCACAATGAGCTCCCACAACACAGCTTTGGTCCACATGATGGTCA 798

Qy   11700 cacttggatatacccagtgtgggtaggggttggggtattgcaggggcctcccaagagtctc 11759
                 ||| |||||||||||||||||||||||||||||||||||||||||| || |||
Db   799 CACTTGGATATACCCAGTGTGGGTAGGTTGGGGTATTGCAGGGGCCTCCCAACAAATCTC 858

Qy   11760 tttaaataataaagaggagttgttcaggt 11787
                 |||||||||||||||||||||||||||||
Db   859 TTTAAATAATAAAGAGGAGTTGTTCAGGT 886

Qy   10983 tggccggccgcccaacaactgcgctccattctctttagagcgccagcgccccgggggcggg 11042
Db   1121 TGGAGCCACCCCACCGCTGCCTCCACCCCTCGAAGTGAGCGCCCCGGGCCCCGGGGGGG 1180

Qy   11043 gtgtgccagccgcgcggggggcggggagccccagcgctcggggcccggttgcggcgcagctcaagcag 11102
Db   1181 GTGTGCGAGGCCGCCGGGGGCCCGGAGCCCAGTGTCGGGGCCGGGTGCGGGGAGCTCAAGCAG 1240

Qy   11103 ttcctcggctggctcaagaagcacgcatactgctcgaaccttagttctccgcctgtacgac 11162
Db   1241 TTCCTCGGCTGGCTCAAGAAGCACCGATACTGCTCGAACCTTAGTTTCCGGCCTGTACGAC 1300

Qy   11163 cagtcggccgtggtcttggatgcagaagtcaacagaccgaaaccagtgtaggaaagttgggg 11222
Db   1301 CAGTGGCGGGTGCTTGGATGCAGAAGTCACACAGACCGAAAACCAGTAGGAAAGTTGGGG 1360

Qy   11223 gaggcgcttgggtgggggttaaagtggaggagagaggaaggaccccgggtgagcagcctcca 11282
Db   1361 GAGGCTTGCGTGCGGGGGTAAAGGAGCAGAGAGGAAGAGACAACCCGGGTGAGCAGCCTCCA 1420

Qy   11283 caacaccgcactctcttcttccaagacacaggacgagagggaaccctcggggcagacggg 11342
Db   1421 CAACACCGCACTCTTCTTCTTCCAAGCACAGGGAGGACGGGGAATCCTGCCCTCGGGGCAGACGGG 1480

Qy   11343 gtgcggcgaggagtaagtggggggtctgggtgaaggtgctcagaggaagtcaga 11402
Db   1481 GTGCGGCGCGGAGGGTAAAGGGGGTCTGGGTGAGTGGGGCCTACAGCAGTCTAGAGATGAGGCCC 1540

Qy   11403 ttcccctcctcggttgttgctcaaagggatctctcattcacccactgcaaa 11462
Db   1541 TTTCCCCTCCTTCGGTTGTTGCTCAAAGGGATCTCTTAGTGTCCTCACCCACTGCAAA 1600

Qy   11463 gagcccaggtttactgcatcatcaagttgctgaagggtccaaggttggcctct 11522
Db   1601 GAGCCCAGGTTTTACTGCATCATCAAGTTGCTGAAAGGGTCCAGGCTTAATGTGGCCTCT 1660

Qy   11523 ttttctgccctcaggttctgtcgcggctaaaactcaaagttgctggtca 11582
Db   1661 TTTCTGCCCTCAGGTCCTGTCGCGGCTAAACTCTAAGGATAGGCCAATCCCTCCTGCTGGGTCA 1720

Qy   11583 gacctgagggcttcaacctgaattggagcccctgtgtacc--atctgggcaacaaagaaacct 11641
Db   1721 GACCTTGGAGGCTCACCTGAAATTGGAGCCCCTTGTGTACTGGGCAACAAATGGGCAAGAAACCT 1780

Qy   11642 acca--gaggctgggcacaatgagctcccacaacacagcttggtccacatgatggtca 11699
Db   1781 ACCATGAGGCTGGGCACAATGAGCTCCCACAACACAGCTTTGGTCCACATGATGGTCA 1840

Qy   11700 cacttggatatacccagtgtgggtaggggttggggtattgcaggggcctcccaagagtctc 11759
Db   1841 CACTTGGATATACCCAGTGTGGGTAGGTTGGGGTATTGCAGGGGCCTCCCAACAAATCTC 1900

Qy   11760 tttaaataataaagaggagttgttcaggt 11787
Db   1901 TTTAAATAATAAAGAGGAGTTGTTCAGGT 1928

RESULT 5
AC003112/c

```
LOCUS       AC003112    40668 bp    DNA    PRI    21-NOV-1997
DEFINITION  Human DNA from chromosome 19 specific cosmid R30292, genomic
            sequence, complete sequence.
ACCESSION   AC003112
NID         g2636669
VERSION     AC003112.1  GI:2636669
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40668)
  AUTHORS   Lamerdin,J.E., McCready,P.M., Adamson,A.W., Burkhart-Schultz,K.,
            Gordon,L., Christensen,M., Kyle,A., Ramirez,M., Stilwagen,S.,
            Garnes,J., Danganan,L., Bruce,R., Quan,G., Montgomery,M., Ow,D.,
            Kobayashi,A., Olsen,A.O. and Carrano,A.V.
  TITLE     Sequence analysis of an ~1 Mb region containing the MEF2B gene in
            19p12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 40668)
  AUTHORS   Lamerdin,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-1997) Human Genome Center, Lawrence Livermore
            National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
FEATURES             Location/Qualifiers
     source          1..40668
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="R30292"
                     /chromosome="19"
                     /map="19p12 between UBA52 and D19S451"
                     /cell_line="5H2-B"
                     /clone_lib="LL19NC03"
                     /note="LL19NC03 cosmid library constructed at LLNL from
                     flow-sorted chromosomes from hybrid 5H2-B, which carries
                     chromosome 19 as its only human chromosome."
     misc_feature    complement(6..62)
                     /note="HERV9 retroviral sequence"
                     /standard_name="endogenous retroviral sequence"
     LTR             225..2516
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 62.000"
     repeat_region   2257..2516
                     /rpt_family="5H2-B"
     repeat_region   8951..9234
                     /rpt_family="Alu"
     misc_feature    complement(7505..7783)
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 71.000"
     repeat_region   complement(5425..5722)
                     /rpt_family="Alu"
     repeat_region   complement(6486..6772)
                     /rpt_family="Alu"
     repeat_region   complement(9358..9535)
                     /rpt_family="Alu"
     repeat_region   complement(9445..9505)
                     /rpt_family="LTR12"
     misc_feature    complement(3269..3544)
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 63.000"
     misc_feature    complement(9629..9672)
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: excellent, score: 75.000"
     repeat_region   complement(9928..10123)
                     /rpt_family="Alu"
     misc_feature    10409..10548
                     /note="DDS similarity to AA047548 zf15e02.r1 Soares fetal
                     heart NbHH19W Homo sapiens cDNA clone 377018 5' (1..138);
                     96% identity.--Other overlapping matches:-(10435..10548)
                     DDS similarity to AA136115 zk90b04.r1 Soares pregnant
                     uterus NbHPU Homo sapiens cDNA clone 490063 5' (1..110);
```

(right column)

```
                     93% identity.--(10466..10548) DDS similarity to AA452628
                     zx33f04.r1 Soares total fetus N62HF8 9w Homo sapiens cDNA
                     clone 788287 5' (1..82); 95% identity.--(10486..10548) DDS
                     NbHH19W Homo sapiens cDNA clone 377018 5' (139..221); 100%
                     similarity to AA009412 ze82h02.r1 Soares fetal heart
                     NbHH19W Homo sapiens cDNA clone 365523 5' (1..61); 97%
                     identity."
     misc_feature    10643..10725
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: excellent, score: 100.000--DDS
                     similarity to AA047548 zf15e02.r1 Soares fetal heart
                     NbHH19W Homo sapiens cDNA clone 377018 5' (139..221); 100%
                     identity.--DDS similarity to AA136115 zk90b04.r1 Soares
                     pregnant uterus NbHPU Homo sapiens cDNA clone 490063 5'
                     (111..192); 99% identity.--DDS similarity to AA452628
                     zx33f04.r1 Soares total fetus N62HF8 9w Homo sapiens cDNA
                     clone 788287 5' (83..165); 100% identity.--DDS similarity
                     to AA009412 ze82h02.r1 Soares fetal heart NbHH19W Homo
                     sapiens cDNA clone 365523 5' (62..144); 100% identity."
     misc_feature    11127..11331
                     /note="DDS similarity to AA136115 zk90b04.r1 Soares
                     pregnant uterus NbHPU Homo sapiens cDNA clone 490063 5'
                     (193..394); 94% identity.--DDS similarity to AA009412
                     ze82h02.r1 Soares fetal heart NbHH19W Homo sapiens cDNA
                     clone 365523 5' (145..346); 96% identity.--DDS similarity
                     to AA009693 ze82h02.s1 Soares fetal heart NbHH19W Homo
                     sapiens cDNA clone 365523 3' (429..307); 95%
                     identity.--(11127..11246) DDS similarity to AA047548
                     zf15e02.r1 Soares fetal heart NbHH19W Homo sapiens cDNA
                     clone 377018 5' (222..341); 97% identity.--(11121..11331)
                     predicted exon, program: grail2exons_human_1.3, frame: 0,
                     quality: good, score: 73.000--(11061..11209) DDS
                     similarity to AA452628 zx33f04.r1 Soares total fetus
                     N62HF8 9w Homo sapiens cDNA clone 788287 5' (167..313);
                     98% identity."
     repeat_region   complement(11869..12161)
                     /rpt_family="Alu"
     misc_feature    12502..12581
                     /note="DDS similarity to AA136115 zk90b04.r1 Soares
                     pregnant uterus NbHPU Homo sapiens cDNA clone 490063 5'
                     (395..477); 92% identity.--DDS similarity to AA009412
                     ze82h02.r1 Soares fetal heart NbHH19W Homo sapiens cDNA
                     clone 365523 5' (347..427); 94% identity.--DDS similarity
                     to AA009693 ze82h02.s1 Soares fetal heart NbHH19W Homo
                     sapiens cDNA clone 365523 3' (306..228); 99%
                     identity.--DDS similarity to AA450010 zx33f04.s1 Soares
                     total fetus N62HF8 9w Homo sapiens cDNA clone 788287 3'
                     (309..229); 99% identity.-"
     repeat_region   complement(12966..13240)
                     /rpt_family="Alu"
     repeat_region   complement(13274..13531)
                     /rpt_family="Alu"
     misc_feature    complement(13756..14029)
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 2, quality: good, score: 61.000--Other overlapping
                     matches:-(13756..13826) DDS similarity to AA009412
                     ze82h02.r1 Soares fetal heart NbHH19W Homo sapiens cDNA
                     clone 365523 5' (428..496); 97% identity.--(13756..13984)
                     DDS similarity to AA009693 ze82h02.s1 Soares fetal heart
                     NbHH19W Homo sapiens cDNA clone 365523 3' (227..1); 98%
                     identity.--(13756..13984) DDS similarity to AA450010
                     zx33f04.s1 Soares total fetus N62HF8 9w Homo sapiens cDNA
                     clone 788287 3' (228..1); 99% identity.-"
     repeat_region   complement(14100..14672)
                     /rpt_family="Alu"
     misc_feature    14885..15317
                     /note="DDS similarity to AA406406 zv11e07.s1 Soares NhHPU
                     S1 Homo sapiens cDNA clone 753348 3' (1..433); Score: 858
                     Identity: 431/433 (99%).--(14884..15237) DDS similarity
                     to W37175 zb21a02.r1 Soares fetal lung NbHH19W Homo
                     sapiens cDNA clone 302666 5' (1..355); 94%
                     identity.--(15227..14885) DDS similarity to AA121532
                     zk89c11.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA
                     clone 490004 3' (342..1); 99% identity."
```

misc_feature

DDS similarity to AA127694 zk89c11.r1 Soares pregnant
uterus NbHPU Homo sapiens cDNA clone 490004 5' (126. .467);
99% identity.--(15227. .14897) DDS similarity to W46604
zc32h10.r1 Soares senescent fibroblasts NbHSF Homo sapiens
cDNA clone 324067 5' (328. .1); 98%
identity.--(15227. .15088) DDS similarity to W46604
zc32h10.s1 Soares senescent fibroblasts NbHSF Homo sapiens
cDNA clone 324067 3' (322. .465); 96% identity."

misc_feature
/note="DDS similarity to AA121532 zk89c11.s1 Soares
pregnant uterus NbHPU Homo sapiens cDNA clone 490004 5'
(389. .343); 99% identity.--DDS similarity to AA127694
zk89c11.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA
clone 490004 5' (77. .125); 90% identity.--(15735. .15713)
DDS similarity to W46603 zc32h10.s1 Soares senescent
fibroblasts NbHSF Homo sapiens cDNA clone 324067 5'
(351. .329); 100% identity.--(15735. .15713) DDS similarity
to W46604 zc32h10.s1 Soares senescent fibroblasts NbHSF
Homo sapiens cDNA clone 324067 3' (299. .321); 100%
identity."

misc_feature
complement(15895. .16082)
/note="DDS similarity to AA121532 zk89c11.s1 Soares
pregnant uterus NbHPU Homo sapiens cDNA clone 490004 3'
(431. .390); 100%
identity.--(15970. .15895) DDS similarity to AA127694
zk89c11.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA
clone 490004 5' (1. .76); 100% identity."

repeat_region
17926. .18190
/rpt_family="Alu"

misc_feature
complement(18270. .18438)
/note="predicted exon, Program: grail2exons_human_1.3,
frame: 0, quality: excellent, score:
100.000--(18438. .18284) DDS similarity to W66776
mel7b11.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus
cDNA clone 387741 5' similar to PIR:B38252 B38252
granulocyte colony-stimulating factor receptor precursor
(157. .1); 82% identity.--(18438. .18406) DDS similarity to
IAA049280 mj45d02.r1 Soares mouse embryo NbME13.5 14.5 Mus
musculus cDNA clone 479043 5' similar to SW:IL6B_MOUSE
Q00560 INTERLEUKIN-6 RECEPTOR BETA PRECURSOR (432. .464);
88% identity."

misc_feature
complement(18540. .18697)
/note="predicted exon, Program: grail2exons_human_1.3,
frame: 1, quality: excellent, score:
100.000--(18540. .18438) DDS similarity to
AA121532 zk89c11.s1 Soares pregnant uterus NbHPU Homo
sapiens cDNA clone 490004 3' (431. .390); 100%
identity.--(18697. .18540) DDS similarity to W66776
mel7b11.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus
cDNA clone 387741 5' similar to
PIR:B38252 B38252 granulocyte colony-stimulating factor
receptor precursor (316. .158); 92% identity.--DDS
similarity to AA049280 mj45d02.r1 Soares mouse embryo
NbME13.5 14.5 Mus musculus cDNA clone 479043 5' similar to
SW:IL6B_MOUSE Q00560 INTERLEUKIN-6 RECEPTOR BETA CHAIN
PRECURSOR (274. .431); 93% identity.--"

Query Match   59.9%;
Best Local Similarity   5.1%;   Score 607.4;   DB 11;   Length 40668;
Matches 1862; Conservative   0; Pred. No. 2.6e-118;   Mismatches 981;   Indels 266;   Gaps 40;

```
Db  29227  CACAGGGGGGACTACAATCTCCCCGTGTCGCCCACCGCAGGAGCTCCCTTCCCGAGGGCCCGGG  29168
Qy      1  gcggcgccgcgtgcagtgattactcaccgcggtggggcaccccaccccggccgctgagtgga   60
           |||||| ||| ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  29407  GTGGCGCGCCGAAGCGATTACTCACCGCGGTGGGGCACCCCACCCCGGCCGCTGAGTGGA  29348

Qy     61  tttttccgtcggggggatggaagaagttaggagagaactctctcgcaccgatgggaacta  120
           |||||||||||| |||| |||||||||| || || ||||||||||||||||||||||||
Db  29347  TTTTTCCGTCGGTGGCGGGGGGTGTGAAATAGGTCCCGGGGAAAAGAGCTTCTTACCGCTCCTCCGT  29288

Qy    121  ggaat-----------------------gcaggggttcggggtcccggttccc  146
                                        ||||| ||||||||||||||||
Db  29287  GGGAATCCCCTCTTACTGCGGTGCGGGACCCAGTTCCCGGCGCAGGGCGTCCGGGTCC  29228

Qy    147  caaaggacac-acctccccataagcccactcataagggctccccctgcacgcgcctccggg  205
           |||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  29227  CACAGGGGGGACTACAATCTCCCCGTGTCGCCCACCGCAGGAG  29168
```

```
Db  29227  CACAGGGGGGACTACAATCTCCCCGTGTCGCCCACCGCAGGAGCTCCCTTCCCGAGGGCCCGGG  29168
Qy    206  acatcccatatccaataccgcagatatgatagttgagaagggaccag---------ag  256
           ||| |||||||||||||||||| |||||||||| || ||| |||||||||||   ||
Db  29167  CGGCCCAGAAGCACACCTCAAAACCATAATCCCGCGCGAAGGGACACGCGCACACCAGGGCGGAAG  29108

Qy    257  gcgggagactccctgccttctggcttctccccccccctgcacgaaacgagactaca  316
           ||| |||||||||| ||| |||||||||||||| ||||||||||||||||||||||
Db  29107  GCGAGGGCGGACTCCTGCCTTCCGTCCCTTTTTCTCTCCCCACCAAAGCCTCGGGCCGAG  29048

Qy    317  gcggtggggagggtggcatgaagcttaggtg----gggatcggtaggagcccatgcacc  372
           ||| |||||||||||||||||||||||||||||   ||||||||| ||||||||||||||
Db  29047  GCTACTCCGGAGAAATCGAAAAAGACGCAGAAGGGCAGAATGGAGAGGAGAGTGGGGGGGCGGACC  28988

Qy    373  cagagaaaaggaactggtggcaacttcca-aactctctgggggaaggaaggggctaaag  431
           |||| ||||||||||||||||||||| ||| |||||||||||||||||||||||||| ||
Db  28987  GAGGGCGGTGTGGGGCGTGGGGTGACTGGAGAACTTTAGGCCCCCTGGGAAGGCTGAGGGGCTGGAGAATGG  28928

Qy    432  aggatgaacggggtcaggtactgctcaatgtgtgttggtcggaccaaagtggggtatggg  491
           ||| ||||||||||||||||||||||||||||||| ||||||||||||||||| ||||||
Db  28927  GGGTAGGCGGGACTCAAGGACACCGCTCGGGTGGGGGGGGGGGGGGGCAGAGCTGGGGTGGGGGGG  28868

Qy    492  ----gccccgtaagagggggcgggggaagtggataggaaggaggatcccggtagactgg-----  543
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  28867  CCGAGACAGAGAAGAATTGGGCCGTGGGGAAGAGGAGAGTGGGGTTGGGGGGACC  28808

Qy    543  -agggatcctggaaaagcaccaaggggctgcgagctagg-------------------aa  581
            ||||||||||||||||||||||||||||||||||||||||                  ||
Db  28807  AAAGGGGCCCACTGAAGAGGGCACTGGGGCTGAGAGGAGGGTGCAGGGAGAGTGGGCAGGGTTA  28748

Qy    582  cccattccggagttaagggtaccaggatcccagatgaggggggtgggaaagcctgggggac  641
           |||||||||||||| ||||| ||||||||||| ||||||||||||||||||||||||||||
Db  28747  TGGAAAACGGGATTGGGGGTGCGGGGGTGCCCAGACT-GGGGGTGGGGAAGCCTGGGACTGGCA  28689

Qy    642  ggaccagagaggaggtcccacgggctgtgggaaagagtggggggggcttcgccgcaggag-701
           |||| |||| ||||||||||| ||||||||||||||||||||||||||| || |||||||
Db  28688  GGGGTTGGAGAGGGAGGGTC-CTGGGGCGGTAGGGGTTGAGTTAGGGGGGCTTTTGGGGCGCG  28630

Qy    702  gatgggacgttcaaggagttgtaactcgggagccccggtgacccgcggtgcc  761
           ||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db  28629  GATAGGGGCCGGTTCTCGGGTGGGGCGGAGGGGAGGGGAGGGGGGGGGGGGTGCCC  28570

Qy    762  gcgggggcggtggggaaggc--cggtgcgggggcccacgatcaaccccccccaggggccggggcc  820
           ||| |||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db  28569  GCGGGGGCGGGGCGGGGGCGGGGGGAGCCCCGGGTGCGGGGCCCCCCCCCGAGGGCGGGGGCC  28510

Qy    821  gggccggggggggggggggcgggaggcgcgaattagcgccttgtcaattcggctg  880
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  28509  GGGGCCGGGGGGGCGGGGGGGCGGGGGGGGGCCAGCGGGCCCATTAGCCGCCTTTGTCAA-TTCGGCTG  28451

Qy    881  cttcagacttggctcgtgtccgcgcccagtgacccgcgtgagg---------931
           |||||||||||| |||||||||||||||||||||||||||||||
Db  28450  CTCAGACTTGCTCTCCGCGCTCCGC-GTCCCGCGCCAGCGACGGGGGGGTGGGGCGCGGGGCCC  28392

Qy    931  -----accggagcccccaatctgcaccccgcagagactcgccccgccccatccgggttgca  986
                |||| |||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db  28391  GCCCTCCCCGCCCCGGCCCCGGGGCCCTGCGCCACCGAGGCGCCCGGGGCGGGGGGCC  28332

Qy    987  gtcaccgccccgttgc-gcgccacccccggtgtcgcccgcgggcgcctcgccccaa  1045
           |||||||||||||||| |||||||||||||||||||||||||||||||| ||||||
Db  28331  CCGCGCCCCCGCCCGGCCCCCCAGCCGCCCCATGCCCGCCGCGGGCGGGGCCCCCA  28272

Qy   1046  tccgcgcgggcggggccgcggcgggccgcgccgtcctcgcgcctcgttgtctctgtc  1105
           ||||||||||| ||||||||||||||||||||| ||||||||| ||||||||||||
Db  28271  TCCCGCGCGGGGCCGCGGCGGGGGCGCCGGTTGCTCGCCCTGCTG------CTGCTGCTCTGGTC  28221

Qy   1106  ctcgggggttgcctcggggcgggatcggggagccgtgagtaccgtgcgcctgccccaacct  1165
           |||||||||||||| ||||||||||||||||| ||||||||||||||||||||||||||||
Db  28220  CTCGGGGGGGGGCGGGGGGACGGGGGACATCAGGAGGCGACGGGAGCGACCGCTGAGTGAGCC--CACTCCTTCC  28163

Qy   1166  cccagggaagccgggggatcggggcgccccgggg-----tagtcgggggggatggaagaaggggc  1225
           ||| |||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db  28162  CCCCTGCCCAGGAGGAAGCGGCGGCCGGGACGACGAGGAGCAGCAGCAGCAGGGGGGCAGCGCGCGGGG  28109
```

```
Qy   1226  gcgagcgcacctggacgtccgggacaaggaaggcg--------gccctcgggg 1274
           |||||  || |||||| ||  || || |||||          ||||| ||||
Db  28108  GCGGGGCCCACCTGGACGGGGAACACGGGAGCGCAGGCGCAGGC gccctgggg 28049

Qy   1275  cgccctcacctgtggggctcatggcacaccccagccttccccaagagtacccgttata 1334
           |||| | ||  |||  ||  ||  ||| |||| | |||||| ||  ||| |||||
Db  28048  CGCCCTCACCTCGGGGCTCATGGCACAGGCGGCCCAGAGTCCACTGTCCCTGCTTCCGC 27989

Qy   1335  catcagaggctcttatctgtatcccctttgcgagggctctggccaagtttgaa 1394
           |||| || |||  || ||  |||| |||| ||| |||||  ||  ||| |||
Db  27988  GCCCGAGGCCCTGGTTACCTGGGGACGGGCCAGCTGGGGGCCAGGCCCCGGGGGC 27929

Qy   1395  ggacacgcagtgtcctgggggaccccctccttcaaggggtgctggggacgttgggggcgcac 1454
           |||| || |||| ||||||| |  ||| |   |||| || ||| ||| || |||||
Db  27928  TGGGGGGATCGGCCCTGAGACCTTGCATCGGACGTGGCGGGGATGAGCCC 27869

Qy   1455  gcctgtgtcttggatatcagagcgaagggaagccttccctggccgggggcgcacgcttg 1514
           |||| | ||| ||  ||| || || || |  ||  |||| || ||||||| || |||
Db  27868  TTCGGTCTGGACTCTGGGCCCTG-GGACTGCAGGGACCCTGGGGGCGCGCACGC 27810

Qy   1515  gtgcgttggttggctgggcaaagtggggctcacaagtgtgatgatcc 1574
           || ||  ||||| | ||  |||||| ||| |  ||||| | |||||
Db  27809  GGTTGCGCGCGCTGGGGGGTCCGGGAGTCCCCCAGCGGG---CCCCCAGGGGGGATGAGCCC 27753

Qy   1575  cgggggagggtggggccgttgatcgttagccgactgaaattcggtga------- 1634
           ||||| | |||||||  ||| |  | || ||| ||||| |||||
Db  27752  CACGGGGAGGGGCGGGGGC-TGCAGGGGGGCGGGGAAATCTTTGTGGGGGGGT 27694

Qy   1635  ccctcgggacttgctcctccgtgggtcgggcgcgcgccctgtaagactcc 1694
           |||| || ||| || || |  ||| || || |||| |||   ||| |||
Db  27693  CCCGGGGGACTTGTGTCTCCACGGGGCCCGGGCCGCGGGGT---TGCGGGACCCG 27637

Qy   1695  atggttggtatcctcgaagtcctcacaacggttcattcaagg 1754
           |||||| | || || || || ||||| | |||| ||||||
Db  27636  GCGCAGGGGGGTCCGGGAGTCCCGCGGCGCTTAGAACCGTTTTCAGCCAGG 27577

Qy   1755  ctgcgctggttgtta------- 1797
           |||| | |||||
Db  27576  CTGCGCTGGGGACAACCCAGCCCTCAGCCGAGAGAAATCTTTGTGAGCGTGAGAGGA 27517

Qy   1797  ---ggagcgagcaagcgtgtccgggcacgcgagccgagacttcattgtctaagggc 1853
              || | ||| ||  ||| |||||| |||| ||| ||| ||||| ||| |||||
Db  27516  GAGGGACGGGGAGCCAGTGAGTGCTCGGGCCACCAGCGGGCCCAA-----ACTTGAGGGGCC 27461

Qy   1854  accaagtggggtcagctgccgagagaatcccactgtgccttg 1913
           |||||||||| | |||  | ||| |||| |||| |||| |||
Db  27460  CACTGGGAACCCGGGCCTGCCAGTGAGTGCCGGGGGACTTCTGGGGCTCA 27401

Qy   1914  agcccc-----atcaccaacgcacacatcccccgccaagatgacgcgtctccacatccaga 1968
           ||||||     |||||||| |||||| ||| |  || ||| || || |||| ||| ||||
Db  27400  GGCACTCTGGGGGGCCCCCATCCAAGAACACCCCCCAAGATACGTCTCCGTATCCAGA 27341

Qy   1969  ccctcctgtggggacacaccccaaccaaagagacccccactggcttatgctccgtcac 2028
           |||| | || ||| ||| ||| || ||| | |||| |||| || || |||||||| |||
Db  27340  CCGGCTGTGGGACACACCAGAGACAC-AAAGAATCACACACTGGCTTGTGGCCGTCAC 27282

Qy   2029  cctgcgcctccga-----cgcgcgtccagtcgcgtattcgtcacacatccgg-- 2081
           |||||  |||||     |||| | |||| | |  |||   ||| |||| ||
Db  27281  CCTCCGCCCTCAAACTCGGCGACACCAGAGACACAGGCGTATTACGCACACATCGGG 27222

Qy   2081  gcgctgccatccatcctcacacacacacacacacacacacac 2140
           |||||| |||||||| ||||||||||||||||||||||||||
Db  27221  GCACACACATTCCATCCCGACACACACACACACACACACAC-- 27173

Qy   2141  acacacagacacgcaacacacacacgcccgcacctccggtggtcgtggtcc 2200
           |||| || || | || |||||||||||| |||||| | || ||||| |||
Db  27172  --CACAGGGGAGGCGGACCCAGGGGCACAGAGGCAGCAGCACTCGGCCAGATGTTCAG 27117

Qy   2201  acatttatttcacggggaggcaaccggggtacgcaatggttgagtgcactgactggagat 2260
           |||   |  |||  ||||||| || ||||| | || ||| || || |||| |||| ||
Db  27116  TCACTCCACGCAC------------------------------------------------ 27103
```

```
Qy   2261  cttcccaccactctcaggtacccggggacacaggcacaacacggcaggggcaccac 2320
           |||    ||   ||||| | |||| |  | ||| |  | ||| | ||||  ||||
Db  27103  ----CCCATCCAGGCGACACGGGGCCATGGGGGTCGGCGCACACACTGCTCCCCAGGC 27048

Qy   2321  gctgcgctgctgctctgggctagtagtcttgtg-cagtttgttccgggggtgtctgtggacg 2379
           |||| |||| |  |||||   | || | |||  ||||| || | ||||| | || |||
Db  27047  TCATCCACCCTGTCCTGGGCTGTTTGCGGGGGGGTGCCTGCCCGTAGGGCGTCCTCGGGGCA 26988

Qy   2380  ccctccccgctctgtcaggggacaggaacctacactcctgcttcgtccgttcctggggc 2439
           || |  | || || ||||| ||| || |  | |||||  ||||| ||| |  ||||
Db  26987  GGGTCCGGGGCACTTCTAGGGGGCCACTTGCTGAAAGCCTCCTGTTTCAGGT 26928

Qy   2440  aggtgatgtggtgacacccggggaccctttccgggggagtggttgttgttctccccag- 2499
           |||| || ||||| || |||||||  ||||| |||||||| ||| ||| | ||||||
Db  26927  .GAG------GGGCTACCTTCGGGCCCCTCCCAGGGCGCCTCCCGTGCCCTGCCCCAAGG 26874

Qy   2499  --tagtttttgaatgcaccaacatagcgctaagcgttttcgtcctgggggcctgtcagaggcaa 2556
             |||||| || | |||||| |  ||| |||| || |||  ||||||| | ||||| |
Db  26873  CAGCGCCTGTGAGTGCCACTGTGAGCCCCAAGCTTTGTTT----TCAGCCAGTAGACCAG 26818

Qy   2557  caggcgaagttggccggagtgggggtgggc-------gtgtgttttcttttgggag 2611
           |||| ||| | ||  ||||| ||||| ||        |||||| |  |||| ||||
Db  26817  CAGGCGAAGGTGTAGGGGGTGGGGGCACGTAAGTGGTGTGTTTTTTTCCTTTAAGGGGGA 26758

Qy   2612  agaaattaaataaagaggttctcacacctctgaaatcctgttt-cttttttaataacgcagcaactccg 2671
           |||||| ||||||||||| | ||| | |||| ||| | ||||   ||||| | ||| |  |||| |
Db  26757  AGAAATTAAATAAAGAGATTCTCACACCTTCGGCAATATATGTTTTCTACTTACGGTGTCTTA 26698

Qy   2672  acacctgaccagccagtcggttgggtcctgaaaagtgttatgtcaggtaccaagtcgggacaggag 2731
           || ||| || || ||   |||  |||  |||||| || || | |||| || | | |||| |||||
Db  26697  GCAACTGACCACCAGGGCGTGGGGTCGT-AAAGTGTCGACAGGAGTACCAGCGCGGACAGGAG 26639

Qy   2732  atgggggcccctggggtatgggctggaggaggccaacttggccttcaggga 2791
           ||||| |||| ||||| |||||||||||| ||| |||| ||||| |||||
Db  26638  ATGGGGACCCCCGGGGTATGGGCTGGAGGAGGCCACCCTGGCCTTCAGAGAA 26579

Qy   2792  tctcacacttctcccctttttaaaacacatggttgtt-cttttttaaataacgcagcaactccg 2850
           |||||||| | ||| ||||||| || |||||| |||   |||||| || || |  ||| |||
Db  26578  TCTCACACTTTTCCCTTTTTAAAACACATGGTCGTTCCTTTTTTAATTAACGCAGTGGCTCCA 26519

Qy   2851  cattgggaaaggggga---aataagcttgtataggccccggggctttgttgggaaagggagggga 2907
           |||||| |||||||    |||| ||||||||| ||||| ||| ||| ||| |||||| ||||
Db  26518  GATTGGGAAAAGAGGAGGAGAAAAAGGCTTTTAAAGGCCCCAGCGTCTGCGGGGAAAAGGGAGGGAG 26459

Qy   2908  ag--agggaagaaaaagaggagggggtgtctcctccaggggttaggggctg 2954
           ||    || || |||||||| || || || |   ||||| ||||||| |||
Db  26458  AGCCAGGGGAGTGAAAAAGGAGGAGCTGTTCAAAAGCCCAACCTAGGGATCTG 26410
```

RESULT 7

```
LOCUS       A70384      1673 bp     DNA         PAT     07-MAY-1999
A70384
DEFINITION  Sequence 14 from Patent WO9811225.
ACCESSION   A70384
NID         g4774664
VERSION     A70384.1   GI:4774664
KEYWORDS    .
SOURCE      unidentified.
ORGANISM    unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1673)
AUTHORS     Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
TITLE       A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
JOURNAL     Patent: WO 9811225-A 19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES             Location/Qualifiers
     source          1..1673
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      344 a    550 c    474 g    305 t
ORIGIN
```

```
Query Match      2.4%;   Score 280.4;  DB 5;  Length 1673;
Best Local Similarity   99.6%;   Pred. No. 3.1e-49;
Matches 281;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy  6330 acacagctgtaatcagcccaggacccccaggacccaccctctccctgcaagcta 6389
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 ACACAGCTGTAATCAGCCCCAGGACCCCCAGGACCCACCCTTCTCACCCCTGCAAGCTA 307

Qy  6390 cctgctctatacatggagacacaccctggggccaccgctgactgactgactctca 6449
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 CCTGCTCTATACATGGAGACACACCCTGGGGCCACCGCTGACTGACTGACTCTCA 367

Qy  6450 atggtcgcgcgcctgccctctgagctgtgtccgcctccttaacacctccacccctggcctg 6509
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ATGGTCGCGCGCCTGCCCTCTGAGCTGTGTCCCGCCTCCTTAACACCTCCACCCCTGGCCTG 427

Qy  6510 ccctggctaaccttaatgggtccaggcaggtcaggacaatcctgttgtcacgccc 6569
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CCCTGGCTAACCTTAATGGGTCCAGGCAGGTCAGGACAATCTGTTGTCACGCCC 487

Qy  6570 gagacggcagcagcattctggctcctcctgcctctatgttggct 6611
         ||||||||||||||||||||||||||||||||||||||||||||
Db   488 GAGACGGCAGCAGCATTCTGGCTCCTCCTGCCTCTATGTTGGCT 529
```

RESULT 8

```
LOCUS       A70388      834 bp      DNA         PAT     07-MAY-1999
A70388
DEFINITION  Sequence 18 from Patent WO9811225.
ACCESSION   A70388
NID         g4774667
VERSION     A70388.1   GI:4774667
KEYWORDS    .
SOURCE      unidentified.
ORGANISM    unidentified
            unclassified.
REFERENCE   1 (bases 1 to 834)
AUTHORS     Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
TITLE       A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
JOURNAL     Patent: WO 9811225-A 19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES             Location/Qualifiers
     source          1..834
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             <1..>834
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAB42575.1"
                     /db_xref="PID:e1433696"
                     /db_xref="PID:g4774668"
                     /db_xref="GI:4774668"
                     /translation="PTLLIGSSLQATCSIHGDPGPATAEGLIWTLNGRRLPSELSRLL
                     NRTSTLALALANLNGSRQQSGDNLVCHARDGSIIAGSCLYVGLLPPEKPFNISCWSRNMK
                     DLTCRWTPGAHGETFLHTNYSLKYKLRWIGQDNTCEEYHTVGPHSCHIPKDLALFTPY
                     EIWEATNRLGSARSDVLILDVLDVVTTDPPDVHVSRVGGLEDQLSVRWVSSPALKD
                     FLPQAKYQIRYREVEDSVDWKVVDDVSNQTSCRLAGLKPGTVIFVQVRCNPFGIYGSKK
                     AG"
BASE COUNT      167 a    274 c    225 g    168 t
ORIGIN
```

```
Query Match      2.2%;   Score 254.4;  DB 5;  Length 834;
Best Local Similarity   99.6%;   Pred. No. 1e-43;
Matches 255;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy  6356 cccacccccttctcatcggctctcccccctccccccgcaagctacctgctatacatggagacacct 6415
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCCACCCTTCTCATCGGCTCTCCCCCTCCCCCGCAAGCTACCTGCTATACATGGAGACACCT 60

Qy  6416 ggggccaccgctgaggggggctctactggactcaagtggtccgcgcctgccctctgagctg 6475
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    61 GGGGCCACCGCTGAGGGGGCTCTACTGGACTCCAAGTGGTCCGCGCCTGCCCTCTGAGCTG 120

Qy  6476 tccggcctcctaacacctccaccctggccctggctaaccttaatgggtccagg 6535
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCCGGCCTCCTAACACCTCCACCCTGGCCCTGGCTAACCTTAATGGGTCCAGG 180

Qy  6536 cagcagtcaggagacacaatctgtgtccacccggagacgccggggctggctggctgactcc 6595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGCAGTCAGGAGAGACAATCTGGTGTGTCACCCGGGGCCCTGGCCGCTGGCTGGCTGCTCC 240

Qy  6596 tgcctctatgttggct 6611
         ||||||||||||||||
Db   241 TGCCTCTATGTTGGCT 256
```

RESULT 9

```
LOCUS       AF059293    1716 bp     mRNA        PRI     02-AUG-1998
AF059293
DEFINITION  Homo sapiens cytokine-like factor-1 precursor (CLF-1) mRNA,
```

us-09-037-657-38.rge   Page 16

```
ACCESSION   AF059293
NID         g3372626
VERSION     AF059293.1  GI:3372626
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1716)
  AUTHORS   Elson,G.C.A., Graber,P., Losberger,P., Herren,S., Gretener,D.,
            Menoud,L.N., Wells,T.N.C., Kosco-Vilbois,M.H. and Gauchat,J.F.
  TITLE     Cytokine Type-I Receptor Family Shares Homology With Members of the
            CLF-1, a Novel Soluble Protein Shares Homology With Members of the
  JOURNAL   J. Immunol. (1998) In press
REFERENCE   2  (bases 1 to 1716)
  AUTHORS   Elson,G.C.A., Graber,P., Losberger,P., Herren,S., Gretener,D.,
            Menoud,L.N., Wells,T.N.C., Kosco-Vilbois,M.H. and Gauchat,J.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-APR-1998) Dept. of Immunology, Serono Pharmaceutical
            Research Institute, 14, Chemin des Aulx, Plan-les-Ouates, GE 1228,
            Switzerland
FEATURES             Location/Qualifiers
     source          1..1716
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19p12"
     gene            1..1716
                     /gene="CLF-1"
     CDS             119..1387
                     /gene="CLF-1"
                     /note="similar to cytokine type-1 receptor family members;
                     similar to the sequence presented in GenBank Accession
                     Number AC003112"
                     /codon_start=1
                     /product="cytokine-like factor-1 precursor"
                     /protein_id="AAC28335.1"
                     /db_xref="PID:g3372627"
                     /db_xref="GI:3372627"
                     /translation="MPAGRGPRAAQSARRPPLPLLLLLCVLGAPRAGSGAHTAVIS
                     PQDFTLLIGSSLLAICSVHGDPPGATAEGLYWTLNGRRLPPELSRVLNASTLALAIAN
                     LNSRSORSDNLVCHARDGSILAGSCLIYGLPPEKPVNISCWSKNMKDLTCRWTPGAH
                     GETFLHTNYSLKYKLRWYGQDNTCEEYHTVGPHSCHIPKDLALFTPYEIWVEARNRLG
                     SARSDVLTLDIDLDVVTTDPPVDHVSRVGGLEDQLSVRWVSPPALKDELFQAKYQIRY
                     RVEDSVDWKVVVDDVSNQSCRLAGLRGTVYFVQVRCNPFGIVGSKKAGIWSEWSHPT
                     AASTPRSERPGPGGGACEPGEPSSGPVRRELKQFLGWLKKHAYCSNLSFRLYDQWR
                     AWMQKSHKTRNQDEGILPSGRRGTARGPAR"
     sig_peptide     119..229
                     /gene="CLF-1"
                     /note="putative"
     mat_peptide     230..1384
                     /gene="CLF-1"
                     /note="putative"
                     /product="cytokine-like factor-1"
BASE COUNT      295 a    623 c    525 g    273 t
ORIGIN
```

```
Query Match       1.9%;   Score 220.2;  DB 11;  Length 1716;
Best Local Similarity   86.5%;  Pred. No. 1.5e-36;
Matches  243;  Conservative   0;  Mismatches  38;  Indels   0;  Gaps   0;

Qy   6330  acacaagctgtaataagcccccagagaccccacccctcctctcatcggtctcctccctgcaagcta  6389
           ||||||  ||  |||||||||||||||||||||||| |||| ||||| ||||||||||||||||||| |
Db    234  ACACAACTGTGATCAGTCCCCAGGATCCCACCAGCTTCATCGGTCCTTCCTCCCTGCTGGCCA  293

Qy   6390  cctgctctatacatggagagacacacctgtgggccacgcgtgagggggctcctactgaccttca  6449
           ||||||||  || |||||| || |||  |||||||||||||| ||||| ||||||| |||||||  ||
Db    294  CCTGCTCAGTGCACGGAGACCACCACCGGCCCACGCGCTTCTACTGACCCTCA  353

Qy   6450  atggtcgccgcctgccctcggtcccgcctcttacacacctccaccctggccctgg  6509
           ||| ||||||||||||||||||| ||||||||||| ||||  |||||| || ||||
```

```
Db    354  ACGGGCGCGCGCCTGCCCCCCTGCCCCGCCTCCCGCCTCCGAGCGCTCCCCGGTGTACTCAGCCTGTCGC  413
Qy   6510  ccctggcttaacccttaatgggtccaggcagcagcagtcaggaagacaatctggtgtgtcacgcc  6569
           ||||||  |||||  ||||| |||||||||||||| |||||||| ||| |||| || ||||||| ||||
Db    414  CCCTGGCCAACCTCAATGGGTCCAGGCAGCAGCAGGCTCGGGGGGACAACCTCGTGTGCCACGCC  473

Qy   6570  gagacgggcagcaattctggctgtccctgctctccttctctatgttggc  6610
           | ||| |||||||||||||||||||||||||| |||| ||||| |||||||
Db    474  GTGACGGCAGCATTCTGGCTGTGGCTCCTGCTTCTATGTTGGC  514
```

```
RESULT 10
LOCUS       A70394                    1391 bp    DNA             PAT  07-MAY-1999
DEFINITION  Sequence 24 from Patent WO9811225.
ACCESSION   A70394
NID         g4774672
VERSION     A70394.1  GI:4774672
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1391)
  AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
            Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
            Kikuchi,Y.
  TITLE     A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
  JOURNAL   Patent: WO 9811225-A 19-MAR-1998;
            NICOLA NICOS ANTONY (AU)
FEATURES             Location/Qualifiers
     source          1..1391
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="unnamed protein product"
                     /codon_start=1
     CDS             <1..1053
                     /protein_id="CAB42576.1"
                     /db_xref="PID:e1433703"
                     /db_xref="PID:g4774673"
                     /db_xref="GI:4774673"
                     /translation="TLNGRRLPPELSRVLNASTLALAIALNLNGSRQRSDNLVCHARD
                     GSILAGSCLIYGLPPEKPVNISCWSKNMKDLTCRWTPGAHGETFLHTNYSLKYKLRWY
                     GQDNTCEEYHTVGPHSCHIPKDLALFTPYEIWVEATNRLGSARSDVLTLDILDVVTTD
                     PPDVHSRVGGLEDQLSVRWVSPPALKDELFQAKYQIRYRVEDSVDWKVVDDVSNQT
                     SCRLAGLKPGTVYFVQVRCNPFGIVGSKKAGIWSEWSHPTAASTPRSERPGPGGGACE
                     PRGGEPSSGPVRRELKQFLGWLKKHAYCSNLSFRLYDQWRAWMQKSHKTRNQHRTRGS
                     CPRADGARREVLPDKL"
BASE COUNT      281 a    459 c    417 g    234 t
ORIGIN
```

```
Query Match       1.6%;   Score 190.8;  DB 5;  Length 1391;
Best Local Similarity   73.5%;  Pred. No. 2.5e-30;
Matches  291;  Conservative   0;  Mismatches  37;  Indels   68;  Gaps   1;

Qy   8604  gtgaccacacgaaccccccacccgacgtgcaagtgagccgcgttgggggcctggaggaccag  8663
           |||||||||||||| ||||||||||||||||||||||| |||||||||||||||||||||||||
Db    469  GTGACCACGGACCCCCCCACCCCCGGACGTGCACGTGAGCCGCGTCGGGGGCCTGGAGGACCAG  528

Qy   8664  ctggagtgttgccgtgggtctcaccaccagctctcaaggagatttccttcctccagccaagtac  8723
           |||| ||| ||||| |||||| ||||||||||||||| |||| ||||| |||||||||||||||  |||
Db    529  CTGAGCGGTGCGGCCGCTGGGTGTGGCCCACCCGCACCCGCACGCACGTGGA  588

Qy   8724  cagatccgctcaccgcggtggaggacagcgtggactggaaggtgcccgtcccggccccggacc  8783
           |||||||||||||||||| || |||||| ||||| ||||||||||||||| ||||| |||||||||
Db    589  CAGATCCGCTACCGACGTGGAGGACAGCGTGGACTGGA---------------  626

Qy   8784  cgccctgaccccgccccgcatctgacctcctcaccgtgcaggtggtggatgacg  8843
           |||||||||||||||||||||||| || |||| ||||||||| |||||||||| |||
Db    626  ------------------------------------AGGTGGTGGACGATG  640

Qy   8844  tcagcaaccagacctcctgccgtcctcgcgggcctgaagcccaccgttacttcgtcc  8903
           ||| |||||||||||||||||||||||||| || |||||||||||||||||||||||| |
```

Db   641 TGAGCAACCAGAGACCTCCTGCGCCTGCCTGGCCTGCCTGGCCTGAACCGTGTACTTCGTGC 700
Qy   8904 aagtggcgttgtaacccattcgggatctatgggtccgaaaaaggcgggaatctggaggcgagt 8963
             ||||||| |||||||||| ||| |||||||||||||| ||||||||||||||||||||| |||
Db   701 AAGTGCGCTGACAACCCTTTGCCATCTATGGCCTCCAAGAAAAGCCGGGGATCTGGAGTGAGT 760

Qy   8964 ggagccaccccaccgcgctcgcctccacccctcgaagtg 8999
             ||||| ||||||||||| ||| ||||||||||||| |||
Db   761 GGAGCCACCCCACAGCGCGCCTCCACTCCCGCAGTG 796

RESULT 11
A70393
LOCUS   A70393   560 bp   DNA   PAT   07-MAY-1999
DEFINITION   Sequence 23 from Patent WO9811225.
ACCESSION   A70393
NID   g4774671
VERSION   A70393.1 GI:4774671
KEYWORDS   .
SOURCE   unidentified.
  ORGANISM   unidentified
               unclassified.
REFERENCE   1 (bases 1 to 560)
  AUTHORS   Nicola,N.A., Fabri,L., Farley,A., Nash,A., Willson,T., Rakar,S.,
               Zhang,J., Alexander,W., Hilton,D.J., Kojima,T., Maeda,M. and
               Kikuchi,Y.
  TITLE   A NOVEL HAEMOPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME
  JOURNAL   Patent: WO 9811225-A 19-MAR-1998;
               NICOLA NICOS ANTONY (AU)
FEATURES   Location/Qualifiers
     source   1..560
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT   113 a   182 c   165 g   100 t
ORIGIN

Query Match   1.2%;   Score 141.2;   DB 5;   Length 560;
Best Local Similarity   92.0%;   Pred. No. 7.9e-20;
Matches 149;   Conservative   0;   Mismatches 13;   Indels   0;   Gaps   0;

Qy   8604 gtgaccacggaccccccaccgagtgcacgtggggccctggggaccag 8663
             |||||||||| ||||||||| ||||| |||||| |||||||||||||
Db   382 GTGACCACGGACCCCCCGACGTGCACCCGCGGCCCTGAGGACCAG 441

Qy   8664 ctgagtgtcgcgctcgggtcctcaccacacagctctcaaggattcctcttccaagccaagtac 8723
             ||||||||| || ||||||| || | ||| |||||||| |||||||||| ||||| ||||
Db   442 CTGAGCGTGCGCGTGGGGTGTCGCCACCGGCCCTCAGGATTTCCTTTTCCAAGCCAAATAC 501

Qy   8724 cagatccggctaccgcgtggaggacagcgtggactggaaggtg 8765
             ||||| || ||||||| || |||||||| |||||||||||||
Db   502 CAGATCCGGCTACCAGTGGAGGACAGTGTGGAATGGAAGGTG 543

RESULT 12
AF060868
LOCUS   AF060868   33613 bp   DNA   ROD   09-DEC-1998
DEFINITION   Mus musculus tumor susceptibility protein 101 (tsg101) gene,
               complete cds.
ACCESSION   AF060868
NID   g3776298
VERSION   AF060868.1 GI:3776298
KEYWORDS   .
SOURCE   house mouse.
  ORGANISM   Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 33613)
  AUTHORS   Wagner,R.U., Dierisseau,P., Rucker,E.B. III, Robinson,G.W. and
               Hennighausen,L.
  TITLE   Genomic architecture and transcriptional activation of the mouse
               and human tumor susceptibility gene TSG101: common types of shorter
               transcripts are true alternative splice variants

JOURNAL   Oncogene 17 (21), 2761-2770 (1998)
MEDLINE   99054675
REFERENCE   2 (bases 1 to 33613)
  AUTHORS   Wagner,R.U. and Hennighausen,L.
  TITLE   Direct Submission
  JOURNAL   Submitted (22-APR-1998) LGP, NIDDK, NIH, Bdg. 10, Rm. 9N115,
               Bethesda, MD 20892, USA
FEATURES   Location/Qualifiers
     source   1..33613
                     /organism="Mus musculus"
                     /strain="129/SvJ"
                     /db_xref="taxon:10090"
     gene   1..33385
                     /gene="tsg101"
     promoter   1..2812
                     /gene="tsg101"
     mRNA   join(2813..2885,9233..9317,11875..11940,13615..13778,
                     15529..15655,19105..19171,21556..21647,30206..30408,
                     31594..31833,33056..33385)
                     /gene="tsg101"
                     /product="tumor susceptibility protein 101"
     CDS   join(2844..2885,9233..9317,11875..11940,13615..13778,
                     15529..15655,19105..19171,21556..21647,30206..30408,
                     31594..31833,33056..33145)
                     /gene="tsg101"
                     /codon_start=1
                     /product="tumor susceptibility protein 101"
                     /protein_id="AAC83576.1"
                     /db_xref="PID:g3776299"
                     /db_xref="GI:3776299"
                     /translation="MAVSESQLKKMSKYKYRDLTVRQTVNVIAMYKDLKPVLDSYVF
                     NDGSSRELVNLTGTIPVRYRGNIYNIPICLWLLDTYPYNPPICFVKPTSSMIKTGKH
                     VDANGKIYLPILHDWKHPRSELLELIQIMIVIFGEEPPVFSRPTVSASYPPTATGTPP
                     NTSYMPGMPSGISAYSGYPNPSGYPGYPCPYPPAGPYPAATSSQYPSQPVTVGPSR
                     RLDQEVAREVDKNIELLKKDEELSSALEKMENQSENNDIDEVIIPTAPLYKQILNLYA
                     EENAIEDTIFLYGEALRRGVIDLDVFLKHVRLLSRKQFQLRALMQKARKIAGLSDLY"
ORIGIN
BASE COUNT   8959 a   6853 c   7411 g   10390 t

Query Match   1.1%;   Score 132;   DB 12;   Length 33613;
Best Local Similarity   76.4%;   Pred. No. 4.1e-18;
Matches 162;   Conservative   0;   Mismatches 50;   Indels   0;   Gaps   0;

Qy   4045 tgtagcccccagctcccctctctcctcctccccccttctcttcttttgcttctgagacttaatt 4104
             ||| | |||||||| |||||||||||||||||||||| |||||||||||||| |||||||||
Db   29004 TGTGAGACACAGTCATTTACTTTACCCCTCCCTTCCTTTTTTTTTTTTTTTTTT 29063

Qy   4105 ttttcttttttcttttggcttttttgagacagggtttctctgtacagccctggctgccct 4105
             ||||||| |||| |||||||||||||||||||||||| |||||||||||||||||||||
Db   29064 TTTTTTTTTTTTTTTTTTTTGATTTTTGAGACAGGGTTTCTCTGTGTAGCCCTGGCTGTCCT 29123

Qy   4165 ggcaactcattctgtagacccagggctagcctcacaaaactacactgcctctgccttt 4224
             ||||||||||||||| |||||||||||||||||||||||||||| |||||||||||
Db   29124 GGAACTCACTCTGTGGACCCAGGCTGGCCTCAAAACTCAGAAAATCCACCTGCCTCTCC 29183

Qy   4225 ccagtgctggcaactaagaatgtgggccaccac 4256
             |||||||| ||||| ||| ||||||||||||||
Db   29184 CAAGTGCTGGGATTAAAGGTGTGCGCCACCAC 29215

RESULT 13
AF139987/c
LOCUS   AF139987   114898 bp   DNA   ROD   03-JUN-1999
DEFINITION   Mus musculus LIM-kinase1 (Limk1) gene, complete cds; Wbscr1
               (Wbscr1) gene, alternative splice products, complete cds; and
               replication factor C, 40kDa subunit (Rfc2) gene, complete cds.
ACCESSION   AF139987
NID   g4972948
VERSION   AF139987.1 GI:4972948
KEYWORDS   .
SOURCE   house mouse.

```
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1  (bases 1 to 114898)
  AUTHORS  Martindale,D.W., Wilson,M.D. and Koop,B.F.
  TITLE    Comparative genomic sequence analysis of the Williams syndrome
           region of human chromosome 7q11.23 containing the LIMK1, WBSCR1,
           and RFC2 genes with its syntenic region in mouse chromosome 5
  JOURNAL  Unpublished
REFERENCE  2  (bases 1 to 114898)
  AUTHORS  Martindale,D.W., Wilson,M.D. and Koop,B.F.
  TITLE    Direct Submission
  JOURNAL  Submitted (01-APR-1999) Department of Biology, Centre for
           Environmental Health, University of Victoria, P.O. Box 3020,
           Victoria, BC V8W 3N5, Canada
FEATURES             Location/Qualifiers
     source          1..114898
                     /organism="Mus musculus"
                     /strain="ES-129/SVJI"
                     /db_xref="taxon:10090"
                     /clone="Genome Systems Inc BAC17511"
     mRNA            join(2873..3130,4731..4827,18321..18459,18691..18800,
                     20947..21153,22176..22281,22391..22557,25458..25641,
                     26178..26264,27155..27286,29309..29368,29613..29678,
                     31468..31624,33507..33562,33821..33978,34086..34248)
                     /gene="Limk1"
                     /product="LIM-kinase1"
     gene            2873..35418
                     /gene="Limk1"
     CDS             join(3076..3130,4731..4827,18321..18459,18691..18800,
                     20947..21153,22176..22281,22391..22557,25458..25641,
                     26178..26264,27155..27286,29309..29368,29613..29678,
                     31468..31624,33507..33562,33821..33978,34086..34248)
                     /gene="Limk1"
                     /codon_start=1
                     /product="LIM-kinase1"
                     /protein_id="AAD34858.1"
                     /db_xref="GI:4972949"
                     /translation="MRLILLCCTWRERMGEGSSELPVCASCGQRIYDGQYLQALNAD
                     WHADCFRCCECSVSLSHQYEKDGQLFCKKDYWARYGESCHGCSEHITKGLVMVAGEL
                     KHPECFICLACGNFIGDGDTYTLVEHKLYCGQCYQTGVTVVPIVIEQQLPDSPGSHLP
                     HRVTLVIPASAHGKRGLSVSIDPPHGPPGCCTEHSHTHRVQGVDPGCVKNSIH
                     VGDRLEINGTPIRNVPLDEIDLIIQEFSRRLQTLEHDPHDSLGHGPVSDPSPLSSP
                     VHPSGQAASSARQKPVLRSCSIDTSPGTSSLASPASQRKDLGRSESSLRVVCRPHRIF
                     RPSDLIHGEVLGKGFCFOAIKVTHRETGEVVMWKELRFEDETGTFLKEVHVMRCLE
                     HPNVLKFIGVLYKDKRLNFITEYIKGGTLRGIIKNMDSQYPWSQRVEFAKDIASGMAY
                     LHSMNIIIHRDLNSHNCLVRENRNWVVADFGLARLMIDEKNQSEDLRSLKKPDRKKRYT
                     VVGNPYWMAPEMINGRSYDEKVDVFSGIVLCEIIGRVNADPDYLPRTMDFGLNVRGF
                     ERFWETRRGESSLPAIPEVPD"
     polyA_signal    35395..35400
                     /gene="Limk1"
     mRNA            join(52154..52241,63700..63887,65870..65934,66013..66109,
                     69389..69526,69827..71599)
                     /gene="Wbscr1"
                     /product="Wbscr1 alternative spliced product"
     mRNA            join(52154..52241,63700..63887,65870..65934,66013..66109,
                     66975..67034,69389..69526,69827..71599)
                     /gene="Wbscr1"
                     /product="Wbscr1"
     gene            52154..71599
                     /gene="Wbscr1"
     CDS             join(52183..52241,63700..63887,65870..65934,66013..66109,
                     66975..67034,69389..69526,69827..69966)
                     /gene="Wbscr1"
                     /codon_start=1
                     /product="Wbscr1"
                     /protein_id="AAD34859.1"
                     /db_xref="PID:g4972950"
                     /db_xref="GI:4972950"
                     /translation="MADDTYDDRAYSSFGGGRGSRGSAGGHGSRSROKELPTEPPYTA
                     YVGNLPFNTVQGDIDAIFKDLSIRSVRLVRDKDTDKFKGFCYVEFDEVDSLKEALTYD
```

CDS
```
                     GAILGDRSLRVDIAEGRKODKGFGEFRKGGPGDRGMGSSRESRGGWDSRDDFNSGYRD
                     DFLGGRGGSRPGDRRAGPPMGSRRDGPPLRGSNMDFREPTEERAQRPRLQIKPRTV
                     ATPLNQVANPNSAIFGGARPREEVVQKEQE"
     CDS             join(52183..52241,63700..63887,65870..65934,66013..66109,
                     69389..69526,69827..69966)
                     /gene="Wbscr1"
                     /codon_start=1
                     /product="Wbscr1 alternative spliced product"
                     /protein_id="AAD34860.1"
                     /db_xref="PID:g4972951"
                     /db_xref="GI:4972951"
                     /translation="MADDTYDDRAYSSFGGGRGSRGSAGGHGSRSROKELPTEPPYTA
                     YVGNLPFNTVQGDIDAIFKDLSIRSVRLVRDKDTDKFKGFCYVEFDEVDSLKEALTYD
                     GAILGDRSLRVDIAEGRKODKGSGFRKGGPGDRGMGSSRESRGGWDSRDDFNSGYRD
                     GSRPRDGPPLRGSNMDFREPTEERAQRPRLQIKPRTVATPLNQVANPNSAIFGGARP
                     REEVVQKEQE"
     polyA_signal    71576..71581
                     /gene="Wbscr1"
     mRNA            complement(join(93152..93630,94484..94597,96116..96196,
                     97205..97270,98133..98290,100260..100360,100502..100603,
                     102020..102126,102503..102549,104514..104583,
                     108595..108710))
                     /gene="Rfc2"
                     /product="replication factor C, 40kDa subunit"
     gene            complement(93165..93169)
                     /gene="Rfc2"
     polyA_signal    complement(93165..93169)
                     /gene="Rfc2"
     CDS             complement(join(93520..93630,94484..94597,96116..96196,
                     97205..97270,98133..98290,100260..100360,100502..100603,
                     102020..102126,102503..102549,104514..104583,
                     108595..108692))
                     /gene="Rfc2"
                     /codon_start=1
                     /product="replication factor C, 40kDa subunit"
                     /protein_id="AAD34861.1"
                     /db_xref="PID:g4972952"
                     /db_xref="GI:4972952"
                     /translation="MEVQESGCDPSESGAOEPSPVPSKIPAGHYELPWVEKIYRPLKNE
                     IVGNEDYVSRLEYFAREGNVPNIIIAGPPGTGKTTSICLARALLGPALKDAVLENA
                     SNDRGIDVVRNKIKMFAQQKVTLPKGRHKIIIDEADSMTDAQQALRITMEIYSKTT
                     RFALACNASDKIIEPIQSRCAVLRYTKLFTDAQVTRLMNVIEKEKVPYTDDGLEAIIF
                     TAQGDMROALNNLQSTFEGFGYINSENVFKVCDEPHPLLVKEMIOHCVDANIEDAYKI
                     LAHWHLGYSPEDVIGNIFRVCKFFPMAEYLKIEFIKEIGYTHMKVAEGVNSLLQOMAG
                     LLARLCOKMAPVAS"
BASE COUNT  28567 a  28771 c  28573 g  28987 t
ORIGIN
```

Query Match 1.18; Score 127.8; DB 12; Length 114898;
Best Local Similarity 69.3%;
Matches 174; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy  4005  ctctggtaactgtcccagggccctctctgacacacaggttgtagcccagctccctc  4064
          |||      |||  ||| |||||||||    |||  |  ||| ||
Db  43519  CTCCCAGGCGCCTGATATAATACTTGCCTGATATTAAACTCTCTTGCCTCCCCCCT  43460

Qy  4065  tttcctccccccctctcttgcttctgagacttaattttttttttttcttttttggc  4124
          |||      ||||||      |||||  |||||  ||||||||||||||
Db  43459  CTCCTCCCTCCCCCCTTCCTTTTTCTTTTCTTTTCTTTTTTTTTTTTTTTT  43400

Qy  4125  ttttttgagacaagggtttcctctgtacagccctggctgcgccctggccactcattctgtagacca  4184
          ||||||||||                  || |||||  |||| || ||  || |||||||||
Db  43399  TTTTTGAGACAGGGTTTCTCTGTACAGCCCTGGCTGTCCTGGGAACTCACTTGTTGAGACCA  43340

Qy  4185  ggctagctcaaactcaaaactaccctcgcctccagtgctggcactaaagat  4244
          |||||||||  ||| ||| || |||||| |||||||| ||||| |||||||
Db  43339  GGCTGGCCCTCAAAACTCAGAAATCCGACCCTCCCAAGTGCTGGGATTAAAGAT  43280

Qy  4245  gtgggcccaca  4255
          ||||| ||||||
Db  43279  GTGCACCACA  43269
```

```
RESULT 14
LOCUS       AC004093   157684 bp   DNA        ROD       31-JAN-1998
DEFINITION  Mus musculus (129SV) DNA, unmapped BAC 10817, complete sequence.
ACCESSION   AC004093
NID         g2828220
VERSION     AC004093.1  GI:2828220
KEYWORDS    HTG.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 157684)
  AUTHORS   Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
            Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,
            Garnes,J., Danganan,L., Poundstone,P., Christensen,M.,
            Georgescu,A., Avila,J., Liu,S., Bruce,R., Quan,G., Montgomery,M.,
            Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O. and
            Carrano,A.V.
  TITLE     Untitled
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157684)
  AUTHORS   Lamerdin,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-1998) Joint Genome Institute, Lawrence Livermore
            National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
FEATURES             Location/Qualifiers
     source          1..157684
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BAC-10817"
                     /cell_line="CJ7"
                     /lab_host="HS996 (modified DH10B/r)"
                     /note="BAC clone obtained frrom Genome Systems. Library
                     made from cell line CJ7, derived from mouse strain 129SV."
     misc_feature    complement(1..646)
                     /note="BLASTN similarity to U17089 (72..503) / match: 0.93,
                     score: 8.5e-139; database searched; nt; Mus musculus MT
                     transposon-like element, clone MT11."
     repeat_region   649..802
                     /rpt_family="L1"
     repeat_region   676..753
                     /rpt_family="L1PA8"
     misc_feature    1516..1563
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 0, quality: excellent, score: 86.000"
     repeat_region   complement(3180..3208)
                     /rpt_family="(GGA)n"
     repeat_region   complement(4735..4991)
                     /rpt_family="POLY_A"
     misc_feature    complement(5770..6199)
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 52.000"
     repeat_region   complement(9823..9908)
                     /rpt_family="(CA)n"
     misc_feature    10150..10173
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 2, quality: excellent, score: 86.000"
     repeat_region   complement(10838..10877)
                     /rpt_family="At_rich"
     repeat_region   complement(11272..11610)
                     /rpt_family="THE1-INTERNAL"
     misc_feature    9771..9823
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 51.000"
     repeat_region   complement(11688..11820)
                     /rpt_family="At_rich"
     repeat_region   complement(11313..13358)
                     /rpt_family="FLAM_C"
     repeat_region   complement(13865..13969)
                     /rpt_family="At_rich"

     repeat_region   complement(14440..14884)
                     /rpt_family="L1MC4"
     misc_feature    14681..14844
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 1, quality: good, score: 67.000"
     repeat_region   complement(15549..15632)
                     /rpt_family="AluJo"
     repeat_region   complement(16026..16058)
                     /rpt_family="(CAAA)n"
     repeat_region   complement(16233..16292)
                     /rpt_family="(CA)n"
     repeat_region   16292..16419
                     /rpt_family="(GA)n"
     repeat_region   complement(17364..17432)
                     /rpt_family="At_rich"
     repeat_region   18848..18901
                     /rpt_family="(CA)n"
     repeat_region   18906..19029
                     /rpt_family="(TAAA)n"
     repeat_region   19108..19156
                     /rpt_family="AluJb"
     repeat_region   complement(19158..19272)
                     /rpt_family="(TAAAA)n"
     repeat_region   19627..19758
                     /rpt_family="(TAGA)n"
     misc_feature    complement(20957..21319)
                     /note="Ribosomal pseudogene.--DPS similarity to
                     gi|1328881|sp|P12947|RL31_HUMAN 60S RIBOSOMAL PROTEIN L31;
                     gi|713328|pir||R5HU31 ribosomal protein L31 - human;
                     gi|713329|pir||R5RT31 ribosomal protein L31 - rat; gi|36130
                     (X15940) ribosomal protein L31 (AA 1-125) [Homo
                     sapiens];gi|157115 (X04809) ribosomal protein L31 (AA
                     1-125) [Rattus norvegicus]; gi|1632855|gnl|PID|e276436
                     (Z81183) ribosomal protein L31 [Sus scrofa]. Score: 405
                     Identity: 91/125 (728)."
     repeat_region   22316..22474
                     /rpt_family="AluJb"
     repeat_region   23291..23335
                     /rpt_family="Alu"
     misc_feature    23316..23350
                     /note="Unresolved number of A's in homopolymeric A-
                     region."
     repeat_region   23336..23434
                     /rpt_family="POLY_A"
     repeat_region   23728..23867
                     /rpt_family="AluJo"
     repeat_region   23874..23925
                     /rpt_family="(GA)n"
     misc_feature    24688..24798
                     /note="predicted exon, program: grail2exons_human_1.3,
                     frame: 2, quality: good, score: 52.000"
     repeat_region   24883..25000
                     /rpt_family="(GA)n"
     repeat_region   complement(25229..25300)
                     /rpt_family="7SLRNA"
     repeat_region   complement(25355..25449)
                     /rpt_family="7SLRNA"
     repeat_region   complement(25390..25465)
                     /rpt_family="AluSp"
     repeat_region   25477..25637
                     /rpt_family="AluSq/x"
     repeat_region   27395..27440
                     /rpt_family="(CA)n"
     misc_feature    27440..27585
                     /note="Low confidence in sequence within low complexity
                     repeat region."
     repeat_region   27441..27566
                     /rpt_family="(CA)n"
     repeat_region   27598..27742
                     /rpt_family="(CA)n"
     repeat_region   complement(28223..28355)
                     /rpt_family="AluSp"
```

```
repeat_region   complement(30491..30518)
                /rpt_family="AT_rich"
repeat_region   complement(30815..30844)
                /rpt_family="AT_rich"
misc_feature    31335..31401
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 0, quality: good, score: 53.000"
repeat_region   complement(31599..31720)
                /rpt_family="(GA)n"
misc_feature    32391..32519
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 2, quality: good, score: 66.000"
repeat_region   complement(33730..33788)
                /rpt_family="LIMB8"
repeat_region   complement(33815..33874)
                /rpt_family="LIMB8"
repeat_region   complement(34206..34568)
                /rpt_family="LIM4"
repeat_region   34803..34858
                /rpt_family="AluSc"
repeat_region   34827..34923
                /rpt_family="FLAM_A"
repeat_region   complement(34928..35353)
                /rpt_family="L1"
repeat_region   35362..35535
                /rpt_family="LIM4"
misc_feature    35462..35731
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 1, quality: good, score: 74.000"
repeat_region   complement(36417..36506)
                /rpt_family="POLY_A"
repeat_region   36874..36894
                /rpt_family="AT_rich"
repeat_region   complement(37247..37346)
                /rpt_family="(GA)n"
repeat_region   complement(37348..37505)
                /rpt_family="FRAM"
misc_feature    38772..39106
                /note="BLASTN similarity to Y08837 (1..335); database searched: nt; H.sapiens mRNA for
                RAD51-like protein"
                score: 1.1e-82;
misc_feature    38773..39400
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 2, quality: good, score: 65.000"
repeat_region   complement(41066..41217)
                /rpt_family="FRAM"
repeat_region   41459..41479
                /rpt_family="AT_rich"
repeat_region   41769..41823
                /rpt_family="AT_rich"
repeat_region   complement(43283..43334)
                /rpt_family="AluJ/monomer"
repeat_region   complement(43492..43523)
                /rpt_family="(CAAA)n"
misc_feature    44337..44447
                /note="predicted exon, program: grail2exons_human_1.3,
                frame: 2, quality: good, score: 72.000"
```

```
Query Match           1.1%;   Score 127.6; DB 12;  Length 157684;
Best Local Similarity  69.9%;  Pred. No. 2.9e-17;
Matches 172;  Conservative  0;  Mismatches 74;  Indels  0;  Gaps  0;

Qy     4050 cccagctccctcctcctccccccctttcgctctgagacttaatttttt 4109
            |||||||| ||| ||||||||||||||| || |||||||||||||||||
Db   154943 CCCTAATTCTCTTTTTTTTGTGTTTTTGTTTTTGGTTTTTTTTTTTTT 154884

Qy     4110 ctttttcttttggctttttgagacagggtttctctgtacaagccctggcac 4169
             || ||| |||| ||| ||||||||||||||||||||| ||| ||||||
Db   154883 GTTGTGTGTGTGTGTTTTTCGGAGACAGGGTTTCTCTGTATAGCCCCTGGAAC 154824

Qy     4170 tcattcgtagaccagggctagcctcaaacctcacacaacctacctgcctctgcctttccagt 4229
            ||| | ||||||||||||| ||||||||||||||||||||||||||||||||||||| |||
Db   154823 TCACTTGTAGACCAGGCTGCCTCAAACCTCAGAAATCCACCTGCCTCTGCCTCCGAGT 154764
```

```
Qy     4230 gctggcactaaagatgtgggccaccacaacctagttaagtgtttgctgtgtcttat 4289
            |||| || ||||| ||| || ||||  |  ||  ||| |||| ||| | ||| || ||
Db   154763 GCTGGGGATTAAAGGCGTGTGTGTCACCATGCCCGGTCCTAATTTTTATAATGTAACACATTG 154704

Qy     4290 tcctat 4295
            ||| |
Db   154703 TCTTTT 154698

RESULT 15
```

```
LOCUS       AC004155    128026 bp    DNA    ROD    28-MAR-1998
DEFINITION  Mus musculus DNA from BAC 10818 containing the Ercc-4 gene,
            complete sequence.
ACCESSION   AC004155
NID         g2896800
VERSION     AC004155.1   GI:2896800
KEYWORDS    HTG.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 128026)
  AUTHORS   Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
            Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,
            Phan,H., Velasco,N., Garnes,J., Danganan,L., Poundstone,P.,
            Christensen,M., Georgescu,M., Avila,J., Liu,S., Attix,C.,
            Andreise,T., Trankheim,M., Amico-Keller,G., Coefield,J., Duarte,S.,
            Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A.,
            Montgomery,M., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O.
            and Carrano,A.V.
  TITLE     Sequence analysis of a mouse BAC containing the DNA repair gene
            Ercc-4 (XPF)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 128026)
  AUTHORS   Lamerdin,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-1998) Joint Genome Institute, Lawrence Livermore
            National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
FEATURES             Location/Qualifiers
     source          1..128026
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="10818"
                     /cell_line="CJ7"
                     /lab_host="HS996 (modified DH10B/r)"
                     /note="Clone 10818 containing mouse Ercc-4 obtained from
                     Genome Systems. CJ7 cell line is derived from mouse strain
                     129SV."
     repeat_region   23..52
                     /rpt_family="(CAAAA)n"
     repeat_region   complement(489..532)
                     /rpt_family="(CCAA)n"
     repeat_region   complement(544..596)
                     /rpt_family="Alus"
     misc_feature    629..661
                     /note="predicted exon, program: grail2exons_mouse_1.3,
                     frame: 1, quality: good, score: 66.000"
     repeat_region   complement(712..742)
                     /rpt_family="AT_rich"
     repeat_region   complement(1415..1582)
                     /rpt_family="(CA)n"
     repeat_region   3254..3316
                     /rpt_family="AluJ/FLAM"
     repeat_region   3516..3576
                     /rpt_family="(CA)n"
     repeat_region   complement(3882..3912)
                     /rpt_family="(CAA)n"
     repeat_region   4480..4601
                     /rpt_family="(GAAAA)n"
     repeat_region   4706..4824
                     /rpt_family="(GGGA)n"
```

```
repeat_region   4979..5064
                /rpt_family="(CA)n"
repeat_region   complement(5418..5522)
                /rpt_family="MIR"
misc_feature    complement(5696..5807)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 0, quality: good, score: 56.000"
repeat_region   complement(6066..6114)
                /rpt_family="(GA)n"
repeat_region   complement(6114..6222)
                /rpt_family="(CA)n"
repeat_region   6824..6917
                /rpt_family="AluSp"
misc_feature    7497..7624
                /rpt_family="(CATA)n"
misc_feature    7886..7911
                /note="predicted exon,
                frame: 1, quality: good, score: 51.000"
repeat_region   complement(9566..9669)
                /rpt_family="LINE2"
repeat_region   10040..10085
                /rpt_family="(CA)n"
repeat_region   complement(11651..11681)
                /rpt_family="AT_rich"
misc_feature    complement(12832..12989)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 2, quality: excellent, score: 83.000"
repeat_region   complement(13670..13799)
                /rpt_family="MER58A"
repeat_region   complement(13821..14009)
                /rpt_family="(CA)n"
repeat_region   14123..14220
                /rpt_family="(CA)n"
repeat_region   14799..14846
                /rpt_family="(CATA)n"
repeat_region   complement(15264..15390)
                /rpt_family="(CAGA)n"
repeat_region   complement(16003..16119)
                /rpt_family="AluJo/FRAM"
repeat_region   16309..16364
                /rpt_family="(CA)n"
repeat_region   16368..16479
                /rpt_family="(GGA)n"
repeat_region   complement(17195..17244)
                /rpt_family="AluSp"
repeat_region   17246..17281
                /rpt_family="(CA)n"
repeat_region   complement(17448..17470)
                /rpt_family="(GA)n"
repeat_region   complement(19060..19219)
                /rpt_family="AT_rich"
repeat_region   19932..20257
                /rpt_family="FRAM"
misc_feature    complement(20409..20528)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 0, quality: good, score: 65.000"
repeat_region   20805..20883
                /rpt_family="(GGA)n"
repeat_region   21143..21236
                /rpt_family="(CA)n"
misc_feature    complement(21689..21837)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 1, quality: good, score: 53.000"
repeat_region   22204..22264
                /rpt_family="AT_rich"
repeat_region   22456..22588
                /rpt_family="(CA)n"
misc_feature    /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 0, quality: excellent, score: 87.000"
repeat_region   23467..23579
                /rpt_family="(GAAAA)n"
repeat_region   23548..23672
                /rpt_family="(GGGA)n"

repeat_region   23663..23807
                /rpt_family="(GA)n"
repeat_region   23842..24064
                /rpt_family="MLT1E"
repeat_region   25418..25485
                /rpt_family="(GGA)n"
repeat_region   complement(25789..25815)
                /rpt_family="AT_rich"
repeat_region   complement(25863..25937)
                /rpt_family="AluSp"
repeat_region   complement(27004..27134)
                /rpt_family="(TA)n"
repeat_region   27997..28083
                /rpt_family="FLAM_A"
repeat_region   28352..28433
                /rpt_family="(CATA)n"
misc_feature    28762..28910
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 1, quality: excellent, score: 76.000"
misc_feature    complement(29273..29397)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 0, quality: excellent, score: 92.000"
repeat_region   29734..29758
                /rpt_family="(CA)n"
repeat_region   complement(29885..30026)
                /rpt_family="AluJo/FRAM"
repeat_region   complement(30337..30398)
                /rpt_family="(TAGA)n"
repeat_region   complement(32477..32523)
                /rpt_family="(CA)n"
repeat_region   32625..32681
                /rpt_family="tRNA-Cys-TGC"
repeat_region   32766..32828
                /rpt_family="POLY_A"
repeat_region   33036..34140
                /rpt_family="L1MA2"
misc_feature    complement(34829..35395)
                /rpt_family="HERVR"
misc_feature    complement(35279..35563)
                /note="BLASTX similarity to P21414 (817..911); match: 0.5,
                score: 2.3e-178; database searched: nr; POL POLYPROTEIN
                (PROTEASE, REVERSE TRANSCRIPTASE"
misc_feature    complement(35714..35770)
                /note="BLASTX similarity to P21414 (709..727); match:
                0.73, score: 2.3e-178; database searched: nr; POL
                POLYPROTEIN (PROTEASE, REVERSE TRANSCRIPTASE"
repeat_region   complement(35765..35979)
                /rpt_family="HERVHC2"
misc_feature    complement(35985..36156)
                /note="BLASTX similarity to P21414 (545..600); match:
                0.73, score: 2.3e-178; database searched: nr; POL
                POLYPROTEIN (PROTEASE, REVERSE TRANSCRIPTASE"
misc_feature    complement(36549..36695)
                /note="BLASTX similarity to P21414 (354..402); match:
                0.55, score: 2.3e-178; database searched: nr; POL
                POLYPROTEIN (PROTEASE, REVERSE TRANSCRIPTASE"
repeat_region   complement(36728..36837)
                /rpt_family="(GA)n"
misc_feature    complement(37079..37151)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 1, quality: excellent, score: 100.000"
misc_feature    complement(37340..37416)
                /note="predicted exon, program: grail2exons_mouse_1.3,
                frame: 2, quality: good, score: 65.000"
```

```
Query Match          1.18;   Score 127.6;  DB 12;  Length 128026;
Best Local Similarity   74.8%;
Matches   160;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Qy   4077  cttctccttttggctcctgagacttaattcttttttcttttttggcttttttgaagacag  4136
           |||||| || |||| | ||||||||||||||||||||||||| |||||| ||||||| |||||
```

```
Db  19029  CTTTCTTTCTTTCTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCAAGACAG 19088
                  ||||||||||||||||||||||||||||||||||||||
Qy   4137  ggtttctctgtacagccctggctgccctggcactcattctgtagaccaggctagcctcaa 4196

Db  19089  GGTTTCTCTGTATAGCCCTGGAtGTCCCTGGAACTCACTTTGTAGACCAGGCTGGTCTTGA 19148
                  ||||||||||||||||||||||||| || ||||||||| ||||||||
Qy   4197  actcacaaacctacctgcctctgcctttccagtgctggcactaaagatgtgggccaccac 4256

Db  19149  ACTCAGAAATCTGCCCGCCTCTGCCCTCCCAAGTGCTGGGATTAAAGGCGTGTGCCACCAC 19208
                  ||||||| ||| |||||||||
Qy   4257  aactagttaagtgttttgctgtgtctttatt 4290
                  ||| |||| |||| | |||
Db  19209  TGCCAGGCCGAAAAGCATTTTCTTAATTACTGATT 19242
```

Search completed: September 19, 1999, 16:18:57
Job time: 16851 sec

OM nucleic - nucleic search, using sw model

Run on: September 19, 1999, 11:40:06 ; Search time 244.65 Seconds
                                              (without alignments)
                                              12100.032 Million cell updates/sec

                        GenCore version 4.5
                        Copyright (c) 1993 - 1998 Compugen Ltd.

Title: US-09-037-657-38
Perfect score: 11832
Sequence: 1 gcggcgcgctgcaagtgattac.........tatgtgctgggtgtggggga 11832

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database :   N_Geneseq_36:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	11830.4	100.0	11832	1	V27148	Nucleotide sequenc
2	6561.6	55.5	6663	1	V27145	Nucleotide sequenc
3	737.2	6.2	1930	1	V27158	Unspliced murine N
4	737.2	6.2	938	1	V27142	Novel haemopoietin
5	280.4	2.4	1629	1	V27141	Novel haemopoietin
6	280.4	2.4	1673	1	V27140	Novel haemopoietin
7	277.2	2.3	1656	1	V41688	Novel haemopoietin
8	254.0	2.2	834	1	V27143	Nucleotide sequenc
9	248	2.1	1724	1	V70896	cDNA encoding rat
10	221.8	1.9	1690	1	V70894	cDNA encoding huma
11	221.8	1.9	1813	1	V70895	cDNA encoding an a
12	220.2	1.9	1579	1	V41689	Nucleotide sequenc
13	190.8	1.6	1391	1	V27144	Nucleotide sequenc
14	141.0	1.2	560	1	V27159	PCR product for hu
15	121.2	1.0	143	1	V27157	Nucleotide sequenc
16	117.6	1.0	2509	1	V27156	Vascular endotheli
17	111.6	1.0	4698	1	O32370	P1A genomic DNA. N
18	111.6	0.9	4698	1	Q72475	Entire P1A genomic
19	111.6	0.9	4698	1	T01176	P815A antigen prec
20	111.6	0.9	4698	1	T14692	MAGE genomic DNA.
21	108.4	0.9	4105	1	X23316	Mouse 1-alpha-OHas
22	108.4	0.9	2534	1	X23318	Mouse 1-alpha-OHas
23	107.4	0.9	259	1	N90724	Nucleotide sequenc
24	107	0.9	6727	1	V70897	Expressed sequence
25	106.6	0.9	3360	1	T88014	Murine IL-5 cDNA g
26	106.4	0.9	14180	1	V19608	Mus musculus EPCR
27	105.4	0.9	3293	1	V05159	Mouse butyrophilin
28	103.6	0.9	4243	1	Q13855	Murine GCSF recept
29	103.4	0.9	210	1	T34501	RHAMM 1 gene intro
30	100.8	0.9	1371	1	V70899	Expressed sequence
31	100.6	0.9	6259	1	V33755	Mouse neutrophil e
32	100.2	0.8	3892	1	T86366	SM22 protein gene
33	100.2	0.8	2190	1	V34385	Upstream region of
34	100	0.8	1681	1	T86367	SM22 protein gene
35	100	0.8	785	1	T86368	SM22 protein gene
36	100	0.8	2147	1	T86369	SM22 protein gene
37	99.8	0.8	535	1	V09117	5' nucleotide sequ
38	99.6	0.8	3364	1	V07266	Muscle creatine ki
39	99.6	0.8	4989	1	O85962	Mouse ST1L gene. A
40	98.4	0.8	5000	1	V38688	Mus musculus SOCS1
41	97	0.8	1371	1	V33755	Mouse neutrophil e
42	96.6	0.8	1371	1	V33755	Mouse neutrophil e
43	96.2	0.8	3419	1	X28077	HS2ST coding seque
44	95.8	0.8	1355	1	Q22971	Sequence encoding
45	95.2	0.8	277	1	X27006	Differentially exp

ALIGNMENTS

RESULT 1
ID V27148
AC V27148;
DT 29-SEP-1998 (first entry)
DT V27148; standard; DNA; 11832 BP.
DE Nucleotide sequence for murine NR6 containing additional 5n sequence.
KW Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW cell survival; therapeutic; neuronal proliferation; drug screening; ssi
OS Mus sp.
OS Mouse.
PN WO9811225-A2.
PD 19-MAR-1998.
PF 11-SEP-1996; G02479.
PR 11-SEP-1997; AU-002246.
PA (AMRA-/ AMRAD OPERATIONS PTY LTD.
PA (DZIE/) DZIEGLEWSKA H E.
PI Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI Zhang J;
DR WPI; 98-260970/23.
PS Claim 9; Fig 3; 182pp; English.
PT New isolated haemopoietin receptor - used for developing products
PT e.g. neuronal cells
PT The NR6 protein is a novel Haemopoietin receptor (HR). Interaction
CC between the novel HR and a ligand facilitates proliferation,
CC differentiation and survival of a wide variety of cells. The HR and it's
CC derivatives can be used for modulating the activity of the receptors e.g.
CC to regulate development, maintenance or regeneration in an array of
CC different cells and tissues in vitro and in vivo. They can be present in
CC therapeutics used for modulating neuronal proliferation, differentiation
CC and survival. The products can also be used for detection and diagnosis,
CC e.g. for cancers or predisposition to cancers, or for drug screening.
SQ Sequence 11832 BP; 2447 A; 3367 C; 3298 G; 2720 T;

Query Match   100.0%; Score 11830.4; DB 1; Length 11832;
Best Local Similarity   100.0%; Conservative 0; Mismatches 0;
Matches 11831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 gcggcgcgctgcaagtgattactcaccgcgctcgcgccaccccacccgcaccggacccctgcaagtgga 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCGGCGCGCTGCAAGTGATTACTCACCGCGCGCACCCCACCCGCGGCGGGCGCCTCAGTGGA 60

Qy   61 ttttccgtgggggggattgaagaagttaggggagaaactcttctgcacgatgggaacta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTTTTCCGTGGGGGGGATTGAAGAAGTTTAGGGAGAAACTCTTCTGCACCGATGGGAACTA 120

Qy   121 ggaatgcaggggttcggttcccgttcccccaaaggagcacacctcctcccccaaagcccactcat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGAATGCAGGGGTTCGGTTCCCGGTTCCCCAAAGGACACACCTCCTCCCCATRAGCCCACTCAT 180

Qy   181 aaggactccctgcacgcgctccgggaactaccccatatccaataccggcagatatgatagt 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGGGCTCCCTGCACGCGCTCCGGGACACCCCATATATCCAATACCGGCAGATATGATAGT 240

Qy   241 tgagaagaggggaccagagaggccggagactccccctcgcctctctggctttccccccgctgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGAGAAGGGACCACCAGAGCCCGGAGACTCCCTCCCCGCTTCTGGCTTTCCCCCCCTGC 300

Qy   301 acgaacgagactacacagccatggagaggcggaaggcggtggtggggggatcgggtgt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACGAAACGAGACTACACAGCCGTGGAGAGAGGTGGCATGAAGGCCTTAGGGTGGGGATCGGGTAG 360

Qy   361 gaccaatgcacccagagaaaggggactgtggtggcaacttcaaactctgggggaaggaaga 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   361  GACCCATGCACCCAGAGAAAAGGACTGGGTGCAACTTTCAAACTCTCTGGGAAGGAAGA              420

Qy   421  aggggctgaaagaggatgaacgggtccaggtactgctccaatgtgtgtcggggaccaaag             480
          |||||  || || || |||| | ||  ||| ||| |||| ||||  |||||| |||||||
Db   421  AGGGGCTGAAAAGAGGATGAAACGGGCTCAGGTACTGCTCAAATGTGTGTCGCGGACCAAAG           480

Qy   481  tgggtatgggggccccgtaagagggggaagtyggataggaggaggatcccggtagact              540
          |||| || |||||||| ||||| ||  || ||||| ||  |||||||  ||| ||||||
Db   481  TGGGTATGGGGGCCCCGTAAGAGGGGGGGATAGGAGGAGGATCCCGGTAAGAGGATCC             540

Qy   541  ggaggggatcctggaaaagcaccagggctcggagctaggaagccattcggagttaagggt            600
          ||||||| |||||||||| |||||||||||  |||||  |||||| || ||| |||||||
Db   541  GGAGGGGATCCTGGAAAAGCACCAGGGCTCGGAGCTAGGAAGCCATTCGGAGTTAAGGGT           600

Qy   601  acaggatcccagatgagggggtggaagcctgggaagcggggggaggaggtcc              660
          ||||| ||||| ||| |||||||||||||||||||| |||||||| ||||||
Db   601  ACAGGATCCCAGATGAGGGGGTGGGGAGCCTGGGAAGCGGGGGGACCAGAGGAGGTCC           660

Qy   661  cacgggctgggtgggggcttcaggagtg              720
          |||||||| |||||||||||||| ||||
Db   661  CACGGGCTGGTTGGGGGGCTTCGCGCCAGGAGGCGGGGATGGGACGGTTCAGGAGTG           720

Qy   721  gtaactgggcggagggccggccggggcggggggccgcgg              780
          |||| |||||||||||||||||| |||||||||||||||
Db   721  GTAACTGGGCGGAGGGCCGGCCGGGGCGGGGGGCCGCGG                       780

Qy   781  gtgcggggggcccagatcaacccccccccccccaggagtg              840
          |||||||||||||| |||||||||||||||||||||||||
Db   781  GTGCGGGGGGCCCAGATCAACCCCCCCCCCCCCCAGGAGTG                     840

Qy   841  cgggtcgggaggggcgattacgggccttgttcaatttggggccctt              900
          ||||||||| ||||||||||||||||||| ||||||||||||||||
Db   841  CGGGGCGGAGCCGGCGCATTAGGCCCTTGTCAATTCGGGGCCCTT                 900

Qy   901  cgcctgtccgcgcccagtgacgcgcctgaggaccccagagact              960
          ||||||||||||||||||||||||||||||||||||||||||
Db   901  CGCTGTCCGCGCCCAGTGACGCGCCTGAGGACCCCAATCTGCACCCCGCAGACT       960

Qy   961  cgcccccgcccccataccggttgcagtcaccggtcaccccccatgccgc           1020
          |||||||||||||  |||||||||||||||||||||||||||||||||||
Db   961  CGCCCCGCGCCCCACCTACCGGCTTGCAGTCACCGGTCACCCCCCATGCCGC         1020

Qy  1021  gggtcgccggggcccgtcggccaatccggcgggggcccgtcgtcctcgct          1080
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  GGGTCGCCGGGGCCCGTCGGCCAATCGCCGCGGGGGCCCGTCGTCCTCGCT         1080

Qy  1081  gtggtcgccctctgttgctctgtgtcctcggggtgcctcggggatcgggag           1140
          ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  1081  GTGGTCGCCCTCTGTTGCTCTGTGTCCTCGGGGTGCCTCGGGGATCGGGAG        1140

Qy  1141  gtaccgttcagtcaacggttgcgccaccccccaatgccgc                1200
          |||||||||||||||||||||||||||||||||||||||
Db  1141  GTACCGTCGCCCTCCCCACCTCCCCAGGGAAGCCGGGGGGT                  1200

Qy  1201  agtccgggggatggaagagggcacccagagtccgggaacaaaggaa             1260
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AGTCGCGGGGGATGGAAGAGGGCACCCAGGGTCCGGGGAACAAAGGAA          1260

Qy  1261  gggcggcgccctcgggggcgccctcacctgtggggctcatgtgcaccacccagctcccaag   1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GGGCGGCGCCCTCGGGGGCGCCCTCACCTGTGGGGCTCATGTGCACCACCCAGCTCCCAAG    1320

Qy  1321  agtacccgttatcatcagaggccctctatctgtatcccccttggaggctgctctgc        1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  AGTACCCCGGTTATACATCAGAGGCCTCTTAATCGTATCCCCCTTTGCGAGTCGGGGGAC    1380

Qy  1381  aggccaagtttgaaggacatcgcagtgtcctggaccctctccttaggggggtagtgggac    1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  AGGCTCAGTTTGAAGGACATCGCAGTGTCCTGGACCCCCTCCTTAGGGTGCTGGGAC        1440

Qy  1441  gcttcggggggcgcacgcctgtgtctggatatcagacggagacggaaggaagcctccctgc   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GCTTCGGGGGGCGCACGCCTGTGTCTGGATATCAGACGGGAAGGGAAGCTTCCCTGCCCG    1500

Qy  1501  gggccgcacgcttgggtgttcggttggggtccctccccat                1560
          |||||||||||||||||||||||||||||||||||||
Db  1501  GGGCCGCACGCTTGGGTGTCGGTTGGGGTCCCTCCCCAT                    1560

Qy  1561  gaagttgatgatccccggggggggaggggggtggggggtcctgttcccggcggc          1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GAAGTGATGATCCCCGGGGGGGAGGGGGGTTATCGTGGCAAAGTGGGGTCCGCGTGGC     1620

Qy  1621  atgcgccccggcgtcccctcgggggacttgcctctccgttgggttcggggtgagg         1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  ATGCGGGCCCGGCGTCCCTTGGGGACTTGCCTCTCCTCCCACTCGGTGGGGTCGGGG      1680

Qy  1681  ctatagcagactccaatgctttggtatcctcggagaagtcctctccaactggtcagacaac   1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  CTATACCAGACTCCAAGCTTTGGTATCCTCGGAAGTCCTCTCCAACTGGTCTCACAAC      1740

Qy  1741  cggtccattcaggctgcgctgggtgtgagagcctaggcgactgaaatttcggtaggag      1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  CGGTCCATTCAGGCGCTCGGGTTGAGAGCCTCTAGGCGACTGAAATTTCGGTGAGGAG     1800

Qy  1801  cgagtgcaaggcggttccgggggcaccggagccgaccggcccggttcaggagtg          1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAGAGCAAGGCCTGTCCGGGGGCACCGGAGCCTCTGGTTGAGAGCCTCATTGTCTAAGGGGCC  1860

Qy  1861  ggggtcagctgccggagaatcccccactgtcccagaacttcattgtctcaaggggccttga    1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  GGGGGTCAGCTGCCGGAGAATCCCCACTGTCCCAGGAGGAACTCCTGGCCTTGAGCCCTGA   1920

Qy  1921  atcaccccaaccgaacaacatccccgccaggagagacttcaagccccagtctctgggga     1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ATCACCCCAACCGCTACAACATCCCCGCCAGGATCCGACCCCTTCCACAGCCCTCTCTGGGA 1980

Qy  1981  cacaccaaagaacacacacacacacacacacaacaaaagagacttatgtcaacctgccccga  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  CACACCCAAAGACACACACAAAAGAGGCCTTATGTCCGCTCACCCTGCCCTCCGA        2040

Qy  2041  cgcggcgctgcagccagatggcgcgctagctccaatcctct             2100
          |||||||||||||||||||||||||||||||||||||||||
Db  2041  CGCGGCGCTGCAGCCCCAGAGTGGCATTCCAATCCTCT                     2100

Qy  2101  acacacacacacacacacacacacacacacacacacacacac           2160
          |||||||||||||||||||||||||||||||||||||||||
Db  2101  ACACACACACACACACACACACACACACACACACACACACAC                 2160

Qy  2161  acacgcacgcacacacacacacgtcggtcgttggtcccacaatttcacacgggag       2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  CACGCACGCACACACACACACGGGTCGTCCCACACTTATTTCACAGGGGAG        2220

Qy  2221  gcaacaccggggtacgccaatatggttcagtgcactgagagatctttccccacaactctagg  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  GCAACACCGGGGTACGGCATATAGGTTCAGATGCACTGAGAGATCTTTCACACGGGAG      2280

Qy  2281  acccccatccggagacaaggccacccgcacacggtcgcgtgctctgggc           2340
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  ACCCCATCCGGAGACAAGGCCACCCGCACACGGTCGCGTGCTCTGGGC           2340

Qy  2341  tagtagttcttgtcgagttttggacgcccgtctgttggtgacagcctcaggg       2400
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TAGTAGGCTTGTCGAGTTTGTGCAGTTTGTCCGCGGTCTGTGGACGCCCTCGTTCAG    2400

Qy  2401  acaggaacctacactcctgcttcgccaaggggggggcccaggtgaatgggcac       2460
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  ACAGGAACCTACACTCCTGCTTTGCCAAGGGGGGCTGAATGGGGTGACACCCGG      2460

Qy  2461  gaccttccgggggagttggttgctgctccaagcctgggtagtttttgaaatgccacaata   2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GACCTTTCCGGGGGAGTTGGTTGCTGCTCCAAGCCTGGGTAGTTTTGGAATGCCACAATA   2520

Qy  2521  gcgctaaagcttttgtttccgggcgggcctgcagcgcaacaggcaacacgtgggggg      2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  GCGCTAAGCTTTGTTTCCGGGCGGGCCTGCAGCGCAACAGGCAACACGCTGAGTGGGGG    2580
```

Qy 2581 tggccgcgttgtgtttttcttttaagggggagagaaattaaataagagggttcctcaacctc 2640
Db 2581 TGGCCGCGTGTGTGTTTTTCTTTTAAGGGGAGAGAAATTAAATAGAGGGTTCTCACACCTC 2640

Qy 2641 tgcaatctgtttgttacttaacgtgtgtcttaacacctgaccagccagccgggtggttcgta 2700
Db 2641 TGCAATCTGTTTGTTACTTAACGTGTGTCTTAACACCTGACCAGCCGGGTGGGTCGTA   2700

Qy 2701 aaagtgtatgcaggtaccacgcgggacacgggggcccctggggtatggctcgggatg 2760
Db 2701 AAAGTGTATGCAGGTACCACGCGGGACACGAGATGGGGTAGGCTGGGATG   2760

Qy 2761 gaggccaccttccgttccaggggaatccacactgttcccttttcccttctca 2820
Db 2761 GAGGCCACCTTCCGTTCCAGGGGAATCCACACTTTTTCCCTTTTAAAACACATG   2820

Qy 2821 gtgttctttttaataacggcagcaactccgcattgggaaaggggaggatagcttgtata 2880
Db 2821 GTGTTCTTTTTAATAACGGCAGCAACTCCGCATTTGGGAAAGGGGAGGATAAGCTTGTATA 2880

Qy 2881 ggcccccggcttcgtttgtgggaaaggggagggaggggaggggggttgtctccca 2940
Db 2881 GGCCCCCGGCTTCGTTTGTGGGAAAGGGGAGGGAGAGGGGGGTGTCCTCCA   2940

Qy 2941 ggcttaggggggctgtcaagctttcgtctgtgtgttgcccccagtcccc 3000
Db 2941 GGCTTAGGGGGGCTGTCAAGCTGTCTGTGTGTTGCCCATGTGTCACAGTCCCC   3000

Qy 3001 agtggctttggcccattgtttgtgtggaagccaagagggggagtcctgagtcctctactcctg 3060
Db 3001 AGTGGCTTTTGGCCCCATTGTTTGTGTGGAAGCCAAGAGGGGGAGTCCTGAGTCTCTACTCTGG 3060

Qy 3061 tactccaggtcaggctcagtccgacccaagaggggaactgagtcttccctgttttatggag 3120
Db 3061 TACTCCAGGTCAGGCTTCCGACCCAAGAGGGGGGACTGAGTCTTCCCTGTTTTATGGAG   3120

Qy 3121 gaatcaggggaaggggtgactaacagttcctgagtgtacaagctcgctcg 3180
Db 3121 GAATCAGGGAAGGGGTGCCAGGTGGACTACGTTCTGCTAGACCCAGTCGCTCG   3180

Qy 3181 aaggagagaaagcttgggcttgcccccctcccaagccaggtatacaagtcgctaggc 3240
Db 3181 AAGGAGAAAGCTTGGGCTTGCCCCCCTCCCAAGCCAGGTAGACGTCTTAATGGAG   3240

Qy 3241 tagtgtggtaaagggcattactcccacaggccaagagagagcccctcctggc 3300
Db 3241 TAGTGTGGTAAAAGGGCATTACTCCCAGGAGAGAGTCCCTGGC   3300

Qy 3301 cagacaaatgtcgtgggggctgggacctccggggcccgccgtctcctg 3360
Db 3301 CAGACAAATGCTGGGACAGAGATGGTGTGATCATTGCCCAGGAGTGCAGACGTGG   3360

Qy 3361 ggtccgggtcgggcagtgctcctcccacctctgaggggggcgtcccagcaggaagcggt 3420
Db 3361 GGTCCCGGGTCGGGCAGTGCTCTCCCACCTGTCTGAGGGCAGCAGGAAGCGGT   3420

Qy 3421 gggtgtgggccggggtagaacgctgcacgtccaagttcatggcctgatta 3480
Db 3421 GGGTGGGCCGGGGTAGAGACGCTGCACGTCCCAGTTCATGCCCAAGTTGAATTA   3480

Qy 3481 gcggggggctgggctgctcggggggcccccgccgtctcctg 3540
Db 3481 GCGGGGGGCTGGGACCTCCCGGGGCGGCCGTCCGTCTG   3540

Qy 3541 gcctgctcctcctgctctcttcgcaaggaaggctgagccctggacaagc 3600
Db 3541 GCCTGCTCCTCCTGCTCTTCGCAGACTTGAGACCCTGGGACAAGC   3600

Qy 3601 cccaaatgcaactgcgattcgcaaggaccgcctcctcccaaggccaaattg 3660
Db 3601 CCCAAAATGCAACTGCGATTGCAAGGAGGATTGAGGGTCACAGGTAGGTTCACAGTCGCC 3660

Qy 3661 cctgggagaagtcattcaaggcccagaactacatgttggtgcacctcaactctcatccatctg 3720

Db 3661 CCTGGGGAGAAGTCATTCAAGGGCCCCAGAACTACATGTTGGTGCACCTCAACTCTCATCCATCTG 3720

Qy 3721 gggcatgaaggaccgtccagggctgcagttagcttcttaatagagaactgggttgggt 3780
Db 3721 GGGCATGAAGGACCGTCCAGGGCTGCAGTTAGCTTCTTAATAGAGAACTGGGTTGGGT 3780

Qy 3781 gcagcctgttcctccgagcctcctttgggaaatcggtttgtttttgtttttcc 3840
Db 3781 GCAGCCTGTTCCTCCGAGCCTCTTTGGGAAATCGGTTTTGTTTTTGTTTTTCC   3840

Qy 3841 aattacctttttcctctcctaacctttcttgcttctacccccaggtgtttctcccccctaagggt 3900
Db 3841 AATTACCTTTTTCCTCTCCTAACCTTTCTTGCTTCTACCCCCAGGTGTTTCTCCCCCCTAAGGGT 3900

Qy 3901 tgcagtcttccctaacctttctttgcttctccacccagtcgtttctcctctttgtttttgttt 3960
Db 3901 TGCAGTCTTCCCTAACCTTTCTTTGCTTCTCCACCCAGTCGTTTCTCCTCTTTGTTTTTGTTT   3960

Qy 3961 cctctccccttgcccaacctttactgactgtgtttctgcttctggttaactgtcc 4020
Db 3961 CCTCTCCCCTTGCCCAACCTTTACTGACTGTTTTCCAGCCTTACTGCCTTGGGTAACGTCC   4020

Qy 4021 caggggcctcctgacacacagggtgtagccccagtccctgctcctctcctctccccttt 4080
Db 4021 CAGGGGCCTCCTGACACACAGGGTGTGAGCCCCAGTCCCTGCTCCTCTCTCTCCCCTTT   4080

Qy 4081 ctcttttgctctgagacttaatttttcttttcttttctcttccctctcctg 4140
Db 4081 CTCTTTGCTTCTGAGACTTAATTTTTTCTTTTCTTTTCTCTTCCCTCTCCTG   4140

Qy 4141 tctctgtacaagcctggactggcccctggcactcaactcctgtagccacctaaaactc 4200
Db 4141 TCTCTGTACAGCCCTGGACTGGCCCCTGGCACTCAATTCTGTAGCCCTACCTAAAACTC   4200

Qy 4201 acaaacctactgcctcttccagttcgttctggtggccactaagaatgttggccacacaact 4260
Db 4201 ACAAACCTACTGCCTCTTCCAGTTCGTTCTGGTGGCCACTAAAGATGTTGGCCACACAACT   4260

Qy 4261 agtagttaagtgtttctgtgtgtcttcatgggcctcagtttcctcagtccctgcatattg 4320
Db 4261 AGTAGTTAAGTGGTTTTTTGCGTGTTCTTTATTCCTAGAGTGCACCTCAGTTTCCTCAGCATAATTG 4320

Qy 4321 taggcgatggatggatggatggatggatgtttgatggagcaag 4380
Db 4321 TAGGCGATGGATGGATGGATGGATGGTTTGATGGAGCCAAGT   4380

Qy 4381 cttgaatcgtcctgagtgaaaaagagaactcagaggaactgaatggagttaggttcccag 4440
Db 4381 CTTGAATCGTCCTGAGTGAAAAAGAGAACTCAGAGGAGTTGAAATGGAGTTAGGTTCCCAG 4440

Qy 4441 ggcagggctggcctggctgtctcatggaggagctccccgtgaaacttccccacacacctcccac 4500
Db 4441 GGCAGGGCCTGGCCTGGCTGTCTCATGGGAGCTCCCCGTGAAACTTCCCCACACACCTCCCAC 4500

Qy 4501 cacctgccatccgtgtggctggccaagagacaatgggacacagactc 4560
Db 4501 CACCCTGCCATCCGTGTGGCTGACAAGGAANTGGGACACAGACTC   4560

Qy 4561 aggaaagcttggaataatgttccctcctcatatcctaggcttgttgtccctgagggc 4620
Db 4561 AGGGAAGCTTGGGAAATAATGTTCCCTCCTCATATCCTAGGCCTTGTTGTCCCTGAGGGC   4620

Qy 4621 ccagcctatgagtaggcagctgggctgcctaagggttgggtaggcaagagaaggggtg 4680
Db 4621 CCAGCCTATGAGTAGGCAGCTGGGCTGCCTAAGGGTTGGGTAGGCAAGAAGGGGTG   4680

Qy 4681 gtccctcaggtgggttcacagtgttcaggttcattccaaaattg 4740
Db 4681 GTCCCTCAGGGTGGGTCACAGTGTTCAGGTTCATTTCCAAAAGTGGCCATCACAGTGCCCT 4740

Qy 4741 aggaaatgattgtggagagtcagaactcctgttgggagttgtagagggccttgcatggtgg 4800

```
Db   4741 AGGAAATGATTGTGGAGAGTCAGAACTCCTGTTGGGGAGTTGTGAGGGGCCTTGCATGTGG 4800

Qy   4801 gcttcctgttgggctgtgcttccttgtcacagtccctgtgttgtgggga              4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GCTTCTGTTGGGCTGTGCTTCCCTTCTCTTGTTGTGGTGTGTTTGTGGGA              4860

Qy   4861 tgtgagggaggggcacgggggaaaatgaaggctcagcccctcagcttgc              4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 TGTGAGGAGGGGCACGGGGAAAAAGAAGGCTCAGCCCCTCAGCTTCACGGTTCAC        4920

Qy   4921 ccaacaggggctcaacctctcctctggacaggcctcactgtagccctcaca            4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 CCAACAGGGGCTCAACCTCTCCTCTGGACAGGCTCACTGTATCCACAGATTGGCCTCACA    4980

Qy   4981 tttgatccctccctttgttgtcctcctgggatgacaggctctcacggttcac           5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 TTTGAATTCCCTTCCTTTGGTCTCTCCACTGGGGAAAAGAAGGCTAGGAGATTTTACA     5040

Qy   5041 tttagatagttccattcctccagaaacacacttgtgaggttaggtatcagttgaaaggac   5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTTTAGATATTGTTCCATTCTTCCAGAAAACACACTTGTGAGGTTAGGTATCAGTGAAAGGAC 5100

Qy   5101 accacaaggacagacaaagaattggaaggaagggaaggtaagccaggccatgtcgttga    5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 ACCACCAGGACAGACAGAAGAATTGGGAGGAGGGAAAATTGGTAAGCCACCGCCATTGCTTGA 5160

Qy   5161 tggcttatgttgtaatcccagaactctgagccgtgaggcaggaggattcccaagtttcaaga  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 TGGCTTATGTGTAATCCCAGAACTCTGGACGCTCGGACAGGAGGATTCCAAGTTCAAGA    5220

Qy   5221 cagtgtgttcctagttaatgagacccgtccaagaaacaactgtgaggaaagaaataaagagaa 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 CAGTGTGTTCTTAGGTAATGAGACCCGTCAAGAAAAGAAAATAAAGAGACAAGAAA      5280

Qy   5281 atgtttattaggtctgtgagacagccttggtctggggcacttggcctccccaatccaagtga 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 ATGTTTATAGGTGGGTGAGACAGCTTGGGGTPAAGGGCCACTGCCTCCAATCAAGATGA    5340

Qy   5341 cctcagccccaccccctagggaatccccatgtgtagaaagacaaactccacgtgtctgacc   5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 CCTCAGCCCCATCCCTAGGAATCTCATGTGTAGAGAAGCAAACTCCAGCGTGGCTAGG     5400

Qy   5401 tccatacactgtgctccaatgtgcacacacaggggagacaataatcaattaatataggag   5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 TCCATACANTGCTCTCCAAATGCACACACAGGGGAGACAATAATCAATTAATAAGGATGTA  5460

Qy   5461 tttgcttagaattgagtaggcattatgactgtgtttaaaaaattttattgatttat      5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 TTTGCTTAGATTTGAGTAGGCANTATGACTGATGTTTAAAATTTTATTGATTTTAT      5520

Qy   5521 gaaaatatacctgttttgtattgtttgagttttgtttattgagacagg             5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 GAAAATATACCTGTTTTGTATTTGGTTTTGTTTATTGAGACAGG                5580

Qy   5581 gcttctctgttctgtgagtctgtggaactcactctgtgagaccaagcaggctgggccttga 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTCTCTGTTGTAGTCTGGGCTGTCTCGTGAGACANTAACTCACTTCGCATGCAGCA    5640

Qy   5641 actcagaaatccgcctgtcttgcttccaagtgctagaattaaggtgttgcactgccat    5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 ACTCAGAAAATCGCCCTGTCTTCCCAAGTGCTTAGATTAAAGGTGTGCACTGCCAT     5700

Qy   5701 tcagcaaaattgcatactttaccccagtatttgggaggcagaggcaggcagactaatgtgtgtga 5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 TCAGCAAAATTGCATACTTTAACCCCAGTATTTGGGGAGGCAGAGGCAGGCAGACTAATGTGTGA 5760

Qy   5761 attccaggctagccaaggatacagagtgagaccctatctcttaccctccccaaaaac      5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 ATTCCAGGCTAGCCAAGGATACAGAGTGAGACCCTATTCTTACCCTCCCCAAAAAC      5820

Qy   5821 ccaaaatgtatttgttcagcacgtaaatgttccaagg                        5880
          ||||||||||||||||||||||||||||||||||||||
Db   5821 CCAAAAGTGTATTTTGTTGTTGGTTGTGTATGTACACGTAAAATGTCCAAGG          5880
```

```
Qy   5881 acaacttgtagaagttctctccgttcacagtccgttccacagtctaagtcctc          5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 ACAACTTGTAGAAGTTCTCTCCGTTCACAGTTCAAGTCTAAGTCCTC               5940

Qy   5941 aggcttagccacagtctctcttcttatgtactgagcacattcactggattgactgat      6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 AGGCTTAGCACACTCTCTTCTTAAGTACTGAGGCCATTTCACTGGCCTGTGATTGACTGAT  6000

Qy   6001 gaattaatttttaagataagtctctttgttgttgttgttcaaacttcaacctccc       6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 GAATTAATTTTTGAGATAAGGTCTAGCTGAGCTAGGCTCAGAAACTATGAACTCCC      6060

Qy   6061 aaggtcatcttgactgtctggttactcttcttctttttcttatctgtaaatgggttg     6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 AAGTCATCTTGACTGTCTGGGTACTCTTCTTCTTTTTTATCTGTTTTATCTGGAATTGAAC 6120

Qy   6121 gcagcacttctctgggatgaggtatctcttggccttggccttcaat              6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GCACCACTTCTCTGGGAAGGGGTGGCCTTGAATTTTGGTTGCCTTCAAT            6180

Qy   6181 gagtgctggggtctggttgttttcttttcttttatctgaaatgggtgaacactgttcaa    6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 GAGTGCTTGGGGTCTGGTTGTTTCGTTGTCTTGTTTATCTGAAATGGGTGAACACTGTTTCAA 6240

Qy   6241 gacttcctgactcttgaaacatccaggcagggtgaggtggctgaagtggtcatccat     6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 GACTTCCTGACTCTTGAAACATCCAGGGCAGGGTGAGGTGGGCTCATCCCAT          6300

Qy   6301 gcctaacaagtgtcgttcttgacccctagcacaagtgataactgagcccccc         6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 GCCTAACAAAGTGTCGTTCTTGACCCCTAGCACAAGTGTAATCCAGGTGTCCCC        6360

Qy   6361 cctttcatcgggctcccctgggccagtgctccctgctaacctttatgcctgat        6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 CCTTCCTAACACTGCCCTGGTCTCCCCGCGAGTCTATACAGGACACCTGTTCAA       6420

Qy   6421 caccgctgagggggtcttactggttacgcctgctgagtctgacctttcccg          6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 CACCGGCTGAGGGGGTCTACTTGAAGTGGGGCTCATCCCAT                    6480

Qy   6481 cctccttaacacccctgggcccctgggccagcatgggcctcaactttagtgggcagca    6540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CCTCCTTAACACCTCACCCTGGGCCCTGGCTAACCTTAATGGGTCCCAGGGACCCCAC    6540

Qy   6541 gtcaggagacaactctgggtgtcacgcccgcagggcactgctgtcctgtcctgcct      6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 GTCAGGAGACAAATCTGGTGTCCAGCCCGAGCCAGCTGTGGGGCCTCTGCCCT        6600

Qy   6601 ctatgttggctgtaagtcgggcccccagaccagagatagggggtcaatgaca         6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 CTATGTTGGCTGTAAGTGGGCCCCAGACCACTGCAGGGGTTGGGCAATGACA         6660

Qy   6661 gattgagcctggggtcttctgtccctgggaggagccatggggctctcactgcatgcagg   6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 GAATTAGAGCCTGGGGTCTTCTGTCCCTGGCAGAGCCATGGGGCTCTCACTTGCATGCAGCA 6720

Qy   6721 catgggtcataccacagcacagcattgcaactcctagggacaggtggctgtgctgtcccc   6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 CATGGTCATACCCACAGCACAGGCATTGCAACTCCTGTTGGGCTACTGTGTCCCC        6780

Qy   6781 tgtgtacccccacagcttagaaaagctgtcaatgtttttcccttgtagttgtgccctgagaagc 6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 TGGGTACCCCACAGCTTTAGAAAGCTGTCATGTTTTCCTTGTGTAGTTGAGTCCCCGAGAAGC 6840

Qy   6841 cctttaacatcagctgtggtccccggaacatgaacatgaaggatctcacgtcaccccacacccg 6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841 CCTTTAACATCAGCTGTGGTCCCGGAACATGAAGGATCTCACGTCCACCCCGGCACACCG    6900

Qy   6901 gtccacagggaggacaattcctcacatctacaagctaagttgg                  6960
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 GTGCACACGGGGAGGACAATTCTTACATACCAACTACTCCCTCAAGTACACAGCTGAGGTTGG 6960
```

```
Qy   6961  tacccagccaagccttgctgtgtgacttcttctctgatacttacttctcctgcaatacttacttctcctgatacaaatg   7020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961  TACCCAGCCAAGCCTTGCCTGTGTGACTTCTGCAATACTTACTTCTCCTGATCAAATATG   7020

Qy   7021  ttcctgtttatgacttcaaaagggacttctgcgcaacctcacaagggtggtacggtcaggtataa   7080
Db   7021  TTCCTGTTTATGAAACTCAAAAAGGGACTCTCGCACCTCCAAGGGTCAGGATAA   7080

Qy   7081  cacactgtgaggagtaccacacactgtgggccctcactaactggaactggc   7140
Db   7081  CACATGTGAGGAGTACCACACACTGTGGGCCCTTGAGGGCAGGAAATGTGTCATCAGAGGACCGC   7140

Qy   7141  cctcttcacttccctatagagatctgagaggcttg   7200
Db   7141  CCTCTTCACTCCCTATAGAGATCTGAGGTGGGAAGGGCCTCACTCATCCCAAGGACCTGGC   7200

Qy   7201  tgatgttcctcaactggatgtctg   7260
Db   7201  TGAATGTCCTCAACTCCCTATGAGATGTGCTCACAGGGTCCTGGAAGGCCTCAGCAAGGACCTC   7260

Qy   7261  gccctagacctttatagggcgctc   7320
Db   7261  GCCCTAGACCTTATAGGGCGCTCCCCCATCCCCAGGACTTTCTTCTTTAGG   7320

Qy   7321  gtcttagccacagggtgtgtttgtcg   7380
Db   7321  GTCTTAGCCACAGGGTGTGTTGTTCATTAACTTAATTCAAAAGC   7380

Qy   7381  tttccccaagagaagtcaagatttt   7440
Db   7381  TTTCCCCAAGAGAAGATTTTTCCCCTCCACACACAACAATACACACA   7440

Qy   7441  cctgcagagaacaacctggcctgaaca   7500
Db   7441  CTCTGCGAGAGAACAACCTGGCTAACACAGTGGTTGTTCATTAAAAGC   7500

Qy   7501  gagtcattagcca   7560
Db   7501  GAGTCCTAGGGACTGAAGGGAGGCGCCCAGGTTCTGAAGGCCGAGGCC   7560

Qy   7561  ttgagctgggtg   7620
Db   7561  TTGAGCTGGGGGGTTGGAGGCACGAGGAGCATGGGTTAATCCTGAGCACAACT   7620

Qy   7621  gggccttaatctaatt   7680
Db   7621  GGGCCTAATCTAATTAGGGTGTTCCCAGCCCATTTAACCCTTCAA   7680

Qy   7681  gtgcctcaactcaagaactcaagg   7740
Db   7681  GTGCCTCACTCAGTGTGGCTCTGCTGCGGGGATCAGCGTTACTCCTCCACGGTCCCCAGGA   7740

Qy   7741  gggttcctggggtcccatt   7800
Db   7741  GGGTTCCTGGGGTCCCCTCATTCCAAGAGGTTTTGGTGCTTTCCTGGCATC   7800

Qy   7801  taacccttagttggtgctcctg   7860
Db   7801  TAACCCTTCAGTTGTTGCTCTGCTGGCACACTGCCCCGTTGGGTTCTTGGTAATGTA   7860

Qy   7861  caaggcatcaaggtggacatg   7920
Db   7861  CAAGGCATCAGAGTGGATGGGGATTACATAGGGATGGAGCCAAATAGCCACTC   7920

Qy   7921  aaggtggggtgatataaat   7980
Db   7921  AAGGTGGGGTGATATAAGCTTGTCACAGGCTCAGAAAGCCTACTCATGAT   7980

Qy   7981  gatcacaattgttgacatcaacctggg   8040
Db   7981  GATCACAATTGTTGACATCAACCTGGGGACACTGTATAGTGAGACCCTAGCTCTCAAAAACACAGAC   8040

Qy   8041  agtagctttaagagtcaagcttgtgacttaatactggaactcaggggcctaataggtgtgctgg   8100
Db   8041  AGTAGCTTTAAGAGTCAAGCAGGTTGAGACTTAATACTGGAACTCAGGGGCCTAATAGGTGCTGG   8100

Db   8101  GTGATGCTCGCCTCACTCCGTTTAGTTGAGATCTCGCTAATCTCCACCCAGCTGG   8160
Qy   8101  gtgtggctgctcctgtttagtgagatctgctaatctccaccca   8160

Db   8161  CTGGGCTGCTCGTCCCCTTGAGGGCAGGGAATGTGTCTTCCATCAGAGATAGGA   8220
Qy   8161  gtgggctgctctgtccccttgagggcaggggaatgtgtcttccatcagagatagga   8220

Qy   8221  tggttgcagtgaactgctgctggtgctgttcctggaattgacagtaatctcatagg   8280
Db   8221  TGGTTGCAGTGAACTGCTGCTGGTGCTGTTCCTGGAATTGACAGTAATCTCATAGG   8280

Qy   8281  ctggggttgagtta   8340
Db   8281  CTGGGGTTGAGTTAGTTAACAGGGGTGGGGCTGGTGTCTGGAAACCGGAGGGTCATAAG   8340

Qy   8341  agcaactgagaccta   8400
Db   8341  AGCAACTGCAGCCTTAGAATTACACACACTGGGGTGTTCGTCACTAGGCCATTCTCACCAAGC   8400

Qy   8401  agtcccaagaactgg   8460
Db   8401  AGTCCCAAGAACTGGGGAGACTGGTTGCCCAGCATTTAATCCCAGCATT   8460

Qy   8461  agggggaggcaagagg   8520
Db   8461  AGGGGAGGGCAAGAGGAGGCGAGAAGGATCTCTGCGAGTTCAAGGCCATCCTGGAGTTCACATAAAG   8520

Qy   8521  agctcccaagcccaagg   8580
Db   8521  AGCTCCCAGGCCCAAGGGTGGGGCGCAGTAAAACCTTGTCTCAAAAAAAAAAACATCTTTAG   8580

Qy   8581  tgacaaggctcctcc   8640
Db   8581  TGACCAGGGTTGCTCCACACCCCAGGGTGCCGCAGTAAAACCTTTAATCCAGCATT   8640

Qy   8641  gcgcgttggggccctggag   8700
Db   8641  GCGGTTGGGGCCTGGAGGGACCACAGCTGAGGTGTGGCCTCCACCACCAGCTCTCAAGG   8700

Qy   8701  atttcctcttcccatcccaagtc   8760
Db   8701  ATTTCCTCTTCCCATCCCAAGTCCGATACCGGTGGGAGGTGGGGCCTGAA   8760

Qy   8761  aggtgcccgtccccgccccggag   8820
Db   8761  AGGTGCCCGTCCCCGCCCCGGATCCCTGACCCCATTATGGGTACGG   8820

Qy   8821  cacgtcaggtggatgacgtaagcaac   8880
Db   8821  CACGTCAGGTGGTGGATGAAACGTCAGGACCAGTGAGTGGCCCTGAA   8880

Qy   8881  gccggtgacacgtttacttcctcc   8940
Db   8881  GCCGGTGACACGTTTACTTCGTCCAAGTGTGTTAAACCCATAGGGGTCAA   8940

Qy   8941  aaggtgccggagaatctggagcgga   9000
Db   8941  AAAGGCGGGAAATCTGGAGCGGAGTGGAGCCACCCACCCACGCCTCCACC   9000

Qy   9001  tgagcaacctcaagggctgg   9060
Db   9001  TGAGCACCTCAGGCTGGTGGAGCGGCATCGGTTCTTCCCTTAAAATGT   9060

Qy   9061  cccacccttttt   9120
Db   9061  CCCACCCTTTTTTGAGACAGCGTCTCAGGTAGGCGCCTTAAATTCAGTATG   9120

Qy   9121  tagtcaaggatgacctcgaagctcctgcttctccacttagagacaaatggccagt   9180
Db   9121  TAGTCAAGGATGACCTCGAAGCTCCTGCTTCTCCACTTAGAGACAAATGGCCAGT   9180
```

```
Db   9121 TAGTCAAGGAGTGAGACCTCGAGCTCCTCGGTCTTTTTGTCTCCACTTAGAGACAATGGCCAGT 9180

Qy   9181 ggccatcacacctttgggaggactggagtctattttgtctgtcatttgtgtgaca 9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9181 GGCCATCACCACCTTTGGGAGACTAGCCATGGAGTCTATTTTAGCCTGTCATTTGGTGACA 9240

Qy   9241 gatggagtacaacagtgtgacctcttgtaagagaagaacttgagagacaggctgtttttaacccc 9300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9241 GATGGAGTACAAACAGTGTGACCTCTTGTAAGAGAGAACTGAGAGACAGGCTGTTTTTAAACCCC 9300

Qy   9301 aatatccctaggctctcctagaggttaacttatatataaaatagagaaactta 9360
          ||||||||||||||||||||||||||||||||||||||||||
Db   9301 AATATCCTAGGCTCTCCTAGAGGTTAACTTTATATATAAAAATAGAGACTATTACACCAGTTA 9360

Qy   9361 tcacatggtcccacagagaacctttgtcacacaacactatagacacaactgcctgtgcctac 9420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9361 TCACATGGTCCCACAGAGAACCTTTGTCACACAACCTTATAGACCACACTAGCCTGTGCCTAC 9420

Qy   9421 cacataaaggtctcctactgtctggcccaaccctcccaacccttaaaaggtaacctaggcagc 9480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9421 CACATAAAGGGTCTCTACTGCTGGCCCACCCCTCCCAACCCTTAAAAGGTAACCTAGGCAGC 9480

Qy   9481 cttaatatttgcaatcctccacttcagcccttccgaatgctcaggaaacaggagtgcattaacc 9540
          ||||||||||||||||||||||||||||||||||||||||||||
Db   9481 CTTAATATTTGCAAATCCTCCTACCCTCTTGAAATGCTCAGGAAACCAGGAGGCATTTAACC 9540

Qy   9541 caagtttctctcttcttcttcggtccccttttctaaagtggggggcctaaagtgacttccttt 9600
          |||||||||||||||||||||||||||||||||||||||||||||
Db   9541 CAAGTTTCTCTCTTCTCTGTGGGTCCCTTTTTCTTAAGGTGGGGAGGGCCTTCTCCC 9600

Qy   9601 gtcctgaagactctccgagccccaaggatctgacactctctctctctaataatgaaataatatt 9660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   9601 GTCCTGAAGACTCTCCGACGCCCATGGATCTGCACACCCCATCCCCACCCGCCTTCCTCCGGTCC 9660

Qy   9661 aatgtctggcctcagtttcccacctcttcttaggtcagcacagagtcggtgtccaaagacac 9720
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   9661 AATGTCTGGCCCTCAGTTTCCCACCTGTGCACCTTTAGGTGGAGCACAGTGTCGGTCCAAGACAC 9720

Qy   9721 ttccttattgcaggcagtatagaagaagaagctcccgagagaatcctcctcctccggttcc 9780
          |||||||||||||||||||||||||||||||||||||||||||||
Db   9721 TTCATTATTTGCAGGCAGGTATAAGAAGAAGCTCCCATCCCCACCCCGCCTTCCTCCGGTCC 9780

Qy   9781 ctaagacagaatacttctacactgaaactgaacctctcgcagacccatatgctcactttaa 9840
          |||||||||||||||||||||||||||||||||||||||||||||
Db   9781 CTAAGACAGAATACTTCTACACTGAAACTGAACTGCGAGACGCGATATGCTCACTTTAA 9840

Qy   9841 tgatgatgaaataatgggagaactaggctccgagagatccctgaggagaaggggtcaaa 9900
          ||||||||||||||||||||||||||||||||||||||||||||
Db   9841 TGATGATGAAATAATGGGAGAACTGAGGCTCCGAGATTCCTGGAGGAGAGGGTCAAA 9900

Qy   9901 accaagctccaggaagctcctccagcacccccatccgggcctctctgggcttctgggcgg 9960
          |||||||||||||||||||||||||||||||||||||||||||
Db   9901 ACCAGCTCCAGGAAGCTCTCCAGCCCCCATCCGGGCCTCCAGTTCTGGGCTTGGGCG 9960

Qy   9961 gagtgaacacagctgggaggaggggagctgggagctttgggcccttgctgccctgccccagc 10020
          |||||||||||||||||||||||||||||||||||||||||||||
Db   9961 GAGTGAACACAGCTGGGAGGGAGGGGGAGCTGGGAGCCTGTTTGGGCCCTTGCTGCCCTGCCCCAGC 10020

Qy  10021 acctcggattcttgcacgtgggggaggggagccagcaggggcgtcgctccggagagactgaagaag 10080
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  10021 ACCTCGCGATTCTTGCCACGTGGGGAGGGGAGCCAGCAGGGCGGCCGGGGAGACTGAAGAAG 10080

Qy  10081 ccgggggttagggttgagggaggtaagcagggggctgttgggggccgaagctctgccaggg 10140
          |||||||||||||||||||||||||||||||||||||||||||
Db  10081 CCGGGGGTAGGGTTGGAGGGGCTGTGGGGCTGTGTGGGGCAGGTGTCCGAAGCTTGTGCCCAGG 10140

Qy  10141 cctgtcagcgagtcccccagtgtttattttattatatgcgtgagtgtcctatccggctgccc 10200
          |||||||||||||||||||||||||||||||||||||||||||
Db  10141 CCTGTCACGCGAGTCCCCCAGTTTATTTATTTGGCTGAGGGTGTCCTTATCGGCTGGCC 10200

Qy  10201 tgctgggggatggctgcggctgggggcgttcccaccacttcactcagtcctc 10260
          ||||||||||||||||||||||||||||||||||||||||||||
Db  10201 TGCTGGGGGATGGCGTCGCGGCTGGGGATGGCGTCGCGGCTGGGGACCCAAGGGCTGCTTCCCACTCCAGTCCTC 10260

Qy  10261 cagcccactcatgtcaacacccgttgcattcctctgaggcttatctctgggaaccgcccttg 10320
          |||||||||||||||||||||||||||||||||||||||||||
Db  10261 CAGCCCACTCCATTGCACACCGTTGCCATTCCTCTGAGCCTTACCTGGGGAACCGCCCTTG 10320

Qy  10321 ttcctgtggctctgtctcctatttctgtcattcctcttccccagagagctgctttttttatggctt 10380
          ||||||||||||||||||||||||||||||||||||||||||||
Db  10321 TTCCTGTGGCTCTGTCTCCTATTTCTGTCATTCACTTCACTGGGATTTTTTCCCAGAGCCTTTTTTTATGCTT 10380

Qy  10381 ttaatataactaacgttttataaaaaattgcttttgtatataatgtttgtgtttctggggtgc 10440
          |||||||||||||||||||||||||||||||||||||||||||
Db  10381 TTAATATAAACTAACGTTTTAAAAATTGCTTTTGTCATTCACTTTCTCCCAGAGCCTTTATGCTT 10440

Qy  10441 gtgccacaacacacgtgaaggttagagaagacttgttgagaagaacttctccaacatgt 10500
          |||||||||||||||||||||||||||||||||||||||||||
Db  10441 GTGCCACAACACACGTGAAGGTTAGAGAGAACTTTGTTGAGTAGGCTCCTCCACCATGT 10500

Qy  10501 gggactagggctggcgacaagagcaattactggggcccctcacccctca 10560
          |||||||||||||||||||||||||||||||||||
Db  10501 GGGACTAGGGCTGGCGACAAGAGCAATTACTGGGCCCCTCACCCCTCA 10560

Qy  10561 cttcccatcctgttctgggatagtcataggttaatcgtaaatacgtgctttaatttcg 10620
          |||||||||||||||||||||||||||||||||||||||||||
Db  10561 CTTCCCATCCTGTTTCTGGGATAGTCATAGGTAAATCGGCTGAAATTCCCCAGCCCCTCA 10620

Qy  10621 tagctatctcctcagcctaccaagtgctctgtcgtaccaagtgctttggggagggctctcc 10680
          |||||||||||||||||||||||||||||||||||||||||||
Db  10621 TAGCTATCCTGCTTTGTCCGTGTCCTACCAAGTGCGTTGTCACCACGTTGTTGGGGAGGGCTCGCC 10680

Qy  10681 tcccagtgtctggggtacacagtccccaagctttcctgcttcctctaggtctcttagtt 10740
          |||||||||||||||||||||||||||||||||||||||||||
Db  10681 TCCCAGTGTCTGGGGTACAGTCCCCAAGTCCCCAAGATCTCCCAAAGGTCCTTTTGTCTTAGTT 10740

Qy  10741 tgtcccttgttctccgttccgtgtcccagagtctccgggccccacttagtccattgattc 10800
          |||||||||||||||||||||||||||||||||||||||||||
Db  10741 TGCCCCTTGCTTTTGTGTCCGTGCCCTGAGGTCTCCCACTTAGTCTCCATTGATTTC 10800

Qy  10801 cttttctgaccgaatactcggttcttacctccccacctgattttgactccctccttgctct 10860
          |||||||||||||||||||||||||||||||||||||||||||
Db  10801 CTTTCTGACCGAATACTCGGTTCTTACCTCCCCACCTGAATTTGACTCCCTCCTTGCTCT 10860

Qy  10861 ccatgccgtggcattcctgggtcctctcatctcctctgggtcccttagtccattttcc 10920
          |||||||||||||||||||||||||||||||||||||||||||
Db  10861 CCATGCCGTGGCAATGCCATTCCCTGGGTGGGTCCCCACTGAGTCCGGGTCCATTAGCCACTTTCC 10920

Qy  10921 caactttcccagccgaagctggttctgttatgggaggcccgtgcccgtcccgcctcct 10980
          |||||||||||||||||||||||||||||||||||||||||||
Db  10921 CAACTTCCCCAGCCGGTCCATCCCCGTGGGTGGAGGCCCGTCCCCGTCCCGCCTCCT 10980

Qy  10981 gctggccgcgcccccaacactgcgctccaacactgtcttagagagcccgggcccggccgcg 11040
          |||||||||||||||||||||||||||||||||||||||||||
Db  10981 GCTGGCCGCGCCCCAACACTGCCGCCTCCAATTCCTTAGAGCGGGCCCCGGGGCCG 11040

Qy  11041 gggtcgtcgcagccgccgcggtggggcgggcggccgcggcagctcaagc 11100
          ||||||||||||||||||||||||||||||||||||||
Db  11041 GGGTGTCCGCGGCGCGGGGCGGGCGGCGCGGCGGCCCGTGGGGCGGGCGGCTCAAGC 11100

Qy  11101 agttcctcggtctggctcaagaagcacgcatactgctcgaacctttcgcctgtacg 11160
          |||||||||||||||||||||||||||||||||||||||
Db  11101 AGTTCCTCGGCTCGGGCTCAAGCAAGCACGCAACACTGCTCGGAACCTTAGTTCGCGGCTGTACG 11160

Qy  11161 accagtgggcgtgcttggagtggatcaacagaacccgaaaccaaggtaggaccaagttgg 11220
          |||||||||||||||||||||||||||||||||||||||||
Db  11161 ACCAGTGGGCGTGGCTTGGATGGCAGAAGCACAGAACCCGAAACCAAGGTAGGAGGCCTC 11220

Qy  11221 gggaggcttgcgtggggggtaaagagagaagagagaagagagcagccctc 11280
          ||||||||||||||||||||||||||||||||||||||||
Db  11221 GGGAGGCTTGCGTGGGGGGTAAAGGAAGGGAGAGAGAGCCCGGTGTGAGGACGCCTC 11280

Qy  11281 cacaacacccgcaactcttcttccaagcacaggacagaggggatcctcggggcagacg 11340
          ||||||||||||||||||||||||||||||||||||||||
Db  11281 CACAACACCCACTCTTCTTTCCAAGCACAGGACACAGGGGATCCCTCGGGGCAGACG 11340
```

```
Qy  11341 gggtgcgcgcgagaggtaagggggtctgggtgagtggggcctacaagcagtctagatgaggc
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11341 GGGTGCGGGCGAGAGGTAAGGGGGTCTGGGTGAGTGGGGCCTACAGCAGTCTAGATGAGGC 11400
Qy  11401 cctttccctccttcgggtgttcgttcgtctcttagtgctcattcacccactgca        11460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11401 CCTTTCCCTCCTTCGGGTGTTCGTTCGTCTCTTAGTGCTCATTCACCCACTGCA 11460
Qy  11461 aagagcccccaggtttactgcatcatcaagttgctgaagggtccaggcttatgtggct
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11461 AAGAGCCCCAGGTTTACTGCATCATCAAGTTGCTGAAGGGTCCAGGCTTAATGTGGCT 11520
Qy  11521 cttttctgccctccaggtctgccggtctaaacctcaagggcatcctcctgctgggt
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11521 CTTTTCTGCCCTCCAGGTCTGCCGGTCTAAACCTCAAGGGCATCCTCCTGCTGGGT 11580
Qy  11581 cagaccttggaggctcacctgattggagcccctcctgtaccactctgggcaacaaagagaaacc
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11581 CAGACCTTGGAGGCTCACCTGATTGGAGCCCCTCCTGTACCACTCTGGGCAACAAAGAGAAACC 11640
Qy  11641 taccagaggctgggcacaatgagctcccacacacagcttggtccacaatgatggttcac
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11641 TACCAGAGGCTGGGCACAATGAGCTCCCACACACAGCTTGGTCCACAATGATGGTTCAC 11700
Qy  11701 acttggatatacccgagtgtgtgggtagggttgggtattgcaggggcctccaaagagtctct
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11701 ACTTGGATATACCCGAGTGTGTGGGTAGGGTTGGGTATTGCAGGGGCCTCCAAAGAGTCTCT 11760
Qy  11761 ttaaataaataaagaagattgttcaagtcccgatggccagtgtgttttggggcctatgtgct
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11761 TTAAATAAATAAAGAAGAGTTGTTCAGGTCCCGATGGCCAGTGTGTTTGGGGCCTATGTGCT 11820
Qy  11821 gggtggggga 11832
           ||||||||||
Db  11821 GGGGTGGGGGA 11832
```

RESULT 2

```
ID    V27145 standard; DNA; 6663 BP.
AC    V27145;
DT    02-OCT-1998 (first entry)
DE    Nucleotide sequence of Murine NR6.
KW    Haemopoietin receptor; cell proliferation; cell differentiation;
KW    cancer; cell survival; therapeutic; neuronal proliferation; drug;
KW    screening; ss; Mouse.
OS    Mus sp.
FH    Key            Location/Qualifiers
FT    CDS            1182..1744
FT                   /*tag=      a
FT                   /product=   "Murine NR6"
FT                   /note=      "No start or stop codon given"
PN    WO9811225-A2.
PD    19-MAR-1998.
PF    11-SEP-1997; G02479.
PR    11-SEP-1996; AU-002246.
PA    (AMRA-) AMRAD OPERATIONS PTY LTD.
PA    (DZIE/) DZIEGLEWSKA H E.
PI    Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI    Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI    Zhang J;
WPI   98-26970/23.
PD    P-PSDB; W55016.
PD    P-PSDB; W55016.
DR    WPI; 98-26970/23.
PT    New isolated haemopoietin receptor - used for developing products
PT    for modulating proliferation, differentiation and survival of cells,
PT    e.g. neuronal cells
PS    Claim 8; Page 108-114; 182bp; English.
CC    The NR6 gene encodes a novel Haemopoietin receptor (HR). Interaction
CC    between the novel HR and a ligand facilitates proliferation,
CC    differentiation and survival of a wide variety of cells. The HR and it's
CC    products can be used for modulating proliferation in an array of
CC    different cells and tissues in vitro and in vivo. They can be present in
```

```
CC    therapeutics used for modulating neuronal proliferation, differentiation
CC    and survival. The products can also be used for detection and diagnosis,
CC    e.g. for cancers or predisposition to cancers, or for drug screening.
SQ    Sequence 6663 BP;
```

```
Query Match   55.5%; Score 6561.6; DB 1; Length 6663;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6628; Conservative 0; Mismatches 19; Indels 5; Gaps 5;

Qy   5186 tggacgctgaggcaggaggagaggatt-ccaagttcaagacagttgtg-ttctaggtaatgagac
          ||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db     12 TGGGACGGCTGAGGCAGGAGGAGGAGGATTCCCAAGTTCAAGACACAGTTGTGTTCTAGGTAATGATGAGAC 71
Qy   5244 cctgtcaagagaaagaaaataaagaggacaagaaatgttatatggtgagagac 5303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 CCTGTCAAGAGAAAGAAAATAAAGAGACAAGAAATGTTATAGGCTGTGAGACAGC 131
Qy   5304 ttgtgtgggtaagggcacttgctcctccaatcaaagatgacctcagcccattcc 5363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 TTGGTGGGTAAGGGCCACTTGCCTCCAATCAAGATGACCTCAGCCCATTCC 191
Qy   5364 catggtagaaggagagacataatcaaaaactc-cagctgctgaccccaatacacatgtctaggaatc 5423
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    192 CATGGTAGAAGGAGAGACATAATCAAAAACTCGACGCTGCTGACCCCAATACACATGTCTCCAAATGTG 251
Qy   5423 cacacacacaggggagacataatcaattaatagatgatgtacttgctagattggactagggca 5482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 CACACACACAGGGGAGACATAATCAATTAATCAATTAATAGATGTATTTGCTTAGATTTGGATAGGGCA 311
Qy   5483 tttatgactgatgtgtttttaaatttttatttgatttttatgtaaatatacctgtttgtattt 5542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 TTTAAGACTGATTGTTTTTAAAATTTTTATTTGATTTTTATGTAAAATATACCTGTTTGTATTT 371
Qy   5543 ggtcttggttgggttttggattttgttttattttgagtttttttt 5602
          |||||||||||||||||||||||||||||||||||||||||||||
Db    372 GGTTTGGTTGGGTTTTGGTTTTGAGTTTTTGTTTTATTTGAGACAGGGCCTTCTCGTGTGTAGTTCCTGGCT 431
Qy   5603 gtccttggaactcactctgtagaccaaggctggcctttgaactcagaaatccgcctgcttgt 5662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GTCCTTGGAACTCACTCTGTAGACCAGGCTGGCCTTGAACTCAGAAATCCGCCTGCTTCTGC 491
Qy   5663 gcttcccaagtgctctagattaaggtgttgcactggcatcagcaaaatgcataactttaa 5722
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GCTTCCCAAGTGCTTAGATTAAAGGTGCACTGGCATCAGCAAAATGCATAACTTTAA 551
Qy   5723 cccagtatttggggagggaggcagaactaaatgtgtgaattccaaggataac 5782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 CCCAGTATTTGGGAGGCAGAGAGACTAAATGTGTGAATTCCAAGGATAAC 611
Qy   5783 agagtgagagacccctattctttaccctccccccccaaaaccccaaaatgtatttgtgctgt 5842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 AGAGTGAGAGACCCCTATTCTTTACCCTCTCCCCCCCAAAACCCCAAAAATGTATTTGTGCTGT 671
Qy   5843 gtatgtacatgtgttgcacgcacgtaaatgtccaaggagcaacttgtagaagttctctcc 5902
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 GTAATGTACAACGTGTTGTTGCACGCACGTAAATGTCCAAGGAGCAACTTGTAGAAGTTCTCTCC 731
Qy   5903 gttcacagtcctaagtcccctgaaatcaaactaaggtcttagcacagtctcttcttt 5962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 GTTCACAGTCCTAAGTCCCTGAAATCAAACTAAGGTCTTAGCACAGTCTCTTCTT 791
Qy   5963 atgttactgagccattcactggcctggattgactgatgaattaattttttgagataaggt 6022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 ATGTACTGAGCCATTCACTGGCCCTGATTGACTGATGAATTAATTTTTGAGATATAGGT 851
Qy   6023 ctccttagctctagctaggctcaaactatgaactccaaggtcatcttgagctgctggt 6082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 CTCTTGTAGCTCTGTAGCTCAGGCTCAAACTATGAACTCCCAAGGTCATCTTGAGCTGCTGGT 911
Qy   6083 actcttgcttccacccacccaagtggttgatgaatactcaggcagcaacttctctggggaggg 6142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 ACTCTTGCTTCCACCCAAGTGGTTGGAATGAATACTCAGGCAGCAACTTCTCTGGGGAGG 971
```

```
Qy  6143 gctggccttggccttgacttgatttgttgctcagcttcaatgagtcgttggtctcgttgtt 6202
         ||||||||||||| ||||| |||||||||||||||||||||||||||| ||||| || ||
Db   972 GCTGGCCCTTGGCCTTGGCCTTGACTTGATTTGTTGTTGCCTCAGCTTCAATGAGTGGCTCGTTGTTT 1031

Qy  6203 cttttctttattcgtgaaatggggtgaacacctgttcaagaacttcaagaacat 6262
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1032 CTTTTCTTTATTCTGTGAAATGGGGTGAACACCTGTTCAAGAACTTCTTGAAACAT 1091

Qy  6263 ccaggcagggtgagggacttgaagtgggctcatcccatgcctaacaaagtgtcgttctttg 6322
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1092 CCAGGCAGGGTGAGGGACTTGAAGTGGGCTCATCCCATGCCTAACAAAGTGTCGTTCTTTG 1151

Qy  6323 acccagacacaagctgtaatcagcccaggaccccaacaccttcctcatcggctcctccctg 6382
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1152 ACCCCAGACACAGCTGTAATCAGCCCAGGACCCCAACACCTTCTCATCGGCTCCTCCCTG 1211

Qy  6383 caagctaccctgctctatacatggaggagacacacccaggagcccactttctactgg 6442
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1212 CAAGCTACCTGCTCTATACATGGAGGAGACACACCCAGGAGCCCACTTTCTACTGG 1271

Qy  6443 accttcaatggtccgccctgagctgttccgctcctccaccctg 6502
         |||||||||||||||||||||||||||||||||||||||||||
Db  1272 ACCTTCAATGGTCCGCCCTGAGCTGTTCCCGCTCCTCCACCCTG 1331

Qy  6503 gccctggcctggctaacttaatgggtccaggaggagaacaatcttggtgtgt 6562
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1332 GCCCTGGCCTGGCTAACCTTAAATGGGTCCAGGAGGAGAACAATCTGGTGTGT 1391

Qy  6563 cacgccgagacgaggcagcgaatatctgttgctcctgtcatatgttgttaagtcgggtgggc 6622
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1392 CACGCCCGAGACGAGGCAGCACATTCTGGGCCTGTCATATGTTGTTAAGTCCACTACTGG 1451

Qy  6623 cccagacaacctcagagatagatggggttggcaatgacacagattctagagcctgggtcttctg 6682
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1452 CCCAGACACTCAGAGATAGATGGGGTTGGCAATGACACAGATTTAGAGCCTGGGTCTGT 1511

Qy  6683 tcctggggcagagccatgggctctcaacttgatgtcatacccagcacaggc 6742
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1512 TCCTGGGGCAGAGCCATGGGCTCTCACTTGGCATAGCCCACAGCCTTACTGG 1571

Qy  6743 attgcaacttaggagacagctgtggtgccactgtcccctatgttggctgtaagtggggc 6802
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1572 ATTGCAACTCTTAGGACAGCCATGGGCTCACTGTCCCATGACCCACAGCCTTTAGAA 1631

Qy  6803 aagctcaatgtttccttgtagtgccccctgagagcccttttacatcagctgctggtc 6862
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1632 AAGCTGTCATGTTTCCTTGTTGTAGTGCCCCTGAGAAGCCCTTTTAACATCAGCTGCTGGTC 1691

Qy  6863 ccggaacatgaaggatctcacgtccgctggacacaggtggcacaggggagacattctt 6922
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1692 CCGGAACATGAAGGANTCACGTGCCGCTGGACACAGGTGGCAGACANTTCTT 1751

Qy  6923 acataccaactactccctcaagtacaagccggaggttgttaccacacaacctgctgctg 6982
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1752 ACATACCAACTACTCCCTCAAGTACAAGCCTAGGTTGGGTACCACCAAGCTTGCTGTG 1811

Qy  6983 tgacttctggcaatacttcctctgatcaaaatatgttcctgttttatgaacttcaaaag 7042
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1812 TGACTTCTGGCAATACTTACCTTCCTCTGATACAAAATAGTTCCTGTTTATGAAACTCAAAAG 1871

Qy  7043 ggactctcgcaacctccaacagtggtacggctcaggaagatacacaagtgaggagtaccacact 7102
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1872 GGACTCTCGCACCTCCACACAGTGGTACGGCTCAGGAATAACACATGGAGGAGTACCACACT 1931

Qy  7103 gtggcccctcactcatgccatatatccccaagggacctccttcctctagagatc 7162
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1932 GTGGGCCCTCACTCATGCCATAATCCCAAGGGACCTCCTTCTCACTCCCTATGAGATC 1991

Qy  7163 tgggtggaaagcccaccacaatcggctcagcaagaatctgatgtgtcctcacaactggatgtc 7222
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1992 TGGGTGGAAAGCCACCACAAATCCCTAGGCTCAGCAAGANTCTGATGTGTCCTCACACTGGATGTC 2051
```

```
Qy  7223 ctggacgtgggtggagccccacagtgtccaactgtgttcctgccctagacctttatagggggcc 7282
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2052 CTGGACGTGGGTGGAGCCCCACAGTGTCCAACCTGTTCCTGGCCCTAGACCTTTATAGGGGGCC 2111

Qy  7283 tccccccaatcccccccagacttttttggtcttctctaggaggtcctcagcagcacggtgg 7342
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2112 TCCCCCCCATCCCCCAGACTTTTTGGTTCTCTCTAGGAGGTCCTCAGCAGCACGGTGG 2171

Qy  7343 ttgcaggacagtggttgtgtttcataactaactttccccccaagacaagtcaaga- 7402
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2172 TTGCAGGACAGTGGTTGTTCATAACTTAATAGTCAAAAGACTTTCCCCCAAGACAAGTCAAGAT 2231

Qy  7402 ttttccctctcccacaccccccaacacacacacacacaccttgcgagaggaaccaccctggcc 7461
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2232 TTTTCCCTCTCCCACCCCCCAACACACACACACACCTTGCGAGAGGAAGACAGACCTGGCC 2291

Qy  7462 tgaccaccttccctctacagccacggtgttcagaaggggactctaggggggactgagagg 7521
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2292 TGACCACCCTTCCCTCTCCTACAGCCCACGGTGTTCAGAAGGGACTCCTAGGGGGACTGAGAGG 2351

Qy  7522 aggcgcccaggtctgaaggcgcccacggaaggccttgaagctgagctggggggggggggcga 7581
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2352 AGGCGCCCAGGTCTGAAGGCGCCCACGGTCAGGAAGCCAGGGACTTGAGCTGGGGGGGGCGA 2411

Qy  7582 gggttggaggcacgaactggttgatcccgtgacacaactgggcctaatctaattaggggtg 7641
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2412 GGGGTTGGAGGCACGAACTCAGCTTGTTACTCTCTCCAAGGTTCCAGGGTTCCGGGGGTG 2471

Qy  7642 ttcccagcccaaaagcagccctgggccattaacccttaaactcaagtgtctgtctgt 7701
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2472 TTCCCAGCCCAAAAGCAGCCCTGGGCCACTTTAACCCTTCAAGTGCTCACTGAAGACTCAAG 2531

Qy  7702 ggagagatcagcttgtactctctccatgttccccagaggaggggttcctgggttgcccctggc 7761
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2532 GGAGGATCAGCTTGTTACTCTCTCCATGGTCCCCAGGAGGGGTTCCTGGGGTTGCCCCTGGC 2591

Qy  7762 tcaattccacaatccacaaggttttgtccttgtcttctcgtcatggatctaaccccagga 7821
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2592 TCATTCCCCACAATCCAGAGGTTTTGTGTCCTTGGTCTTCCGTCANTCTAACCCTCAGGA 2651

Qy  7822 ggctggacagctgccccgtgggaggctcttggcatgtaagaggaagaggttgtgtccttgt 7881
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2652 GGCTGGCACAGCTGCCCCGTGGGAGGCTCTTGGCATGTAAGAGGGCATCAGAGGGTGGGACAT 2711

Qy  7882 gggatggggatacataggggttgagcaaatagcaacctcaggtgggggtgatatacaata 7941
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2712 GGGATGGGGATNCANTAGGGATGAGCCCAAATAGCCACCTCAAGGGTGGGGGTGATATACANTA 2771

Qy  7942 aagcttgtcacctcacgtcagcctcagcagcctttaggatgtgatcatcac 8001
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2772 AAGCTTGTCACCTCACGTCAGCCTCAGCAGCCTTTAGGATGTGATCATCAC 2831

Qy  8002 tcctggggacagttagttgagaaccctagctcaaaacacagatagctttaagagtcagct 8061
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2832 TCTTGGGACAGTTAGTGAGAACCCTAGCTCAAAACACAGATAGCTTTAAGAGTCAGCTT 2891

Qy  8062 gtgacttaatactggaactcaggtgggtgggtgatcgtgccttcaactcact 8121
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2892 GTGACTTAATACTGGAACTCAGGGTGGGTGGGTGATCGTGCCTTCACTYCCT 2951

Qy  8122 gtttagtgagatcctcggcgctaaatccccccaccccctgtggttgcccctg 8181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2952 GTTTAGTGAGATCTCGGCGCTAAATCCCAGGCTGTGGGTGGGTGGGCCCCTG 3011

Qy  8182 agggcaggaatgtgtctccatcagagatagagaccccgtggtagagcaacctgctgctg 8241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3012 AGGGGCAGGAANTGTGTCTTCCATCAGAGATAGGACCCGGTGGTAGGCAACCTGTGCTG 3071

Qy  8242 gctgtttctggaatattaaatgacagtatctctcaggcctgggttggagtagctaacagg 8301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3072 GCTGTTTCTGGAATATTAAATGACAGTAAGCTATCGGCCTGGGTGAGTGAGTAGCTAACAGGG 3131

Qy  8302 gtggggcgtggtctggaaacgcagataggagccactgcagcctagattac 8361
```

```
Db  3132  GTGGGGGGCGTGGGTCGGGAAAACGGACAGAATAGGGTCAGTAGGACGCCACTGGCGCCTAGATTAC  3191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8362  accactgggtgttctgtcactaggccaattccaacaagctgggagcac  8421
          ||||||||||||||||||| |||||||
Db  3192  ACCACTGGGTGTTCTGTCACTAGGCCAATTCTCACCAAGCAGTGCTTCAGAACTGGGAGCAC  3251

Qy  8422  tgttggccagcatttaatgccagcagcattaggggaggggagcaggg caggaagg  8481
          |||||||||||||||||||||||||
Db  3252  TGTTGCCACCATTTAATGCCAGCATTTAATGCCAGCATTAGGGAGGGGCAGAGGCAGAAAGG  3311

Qy  8482  atctctctgagttcaaggccacatctgaattcactacataaagagctccagcagcaggtg  8541
          |||||||||||||||||||||
Db  3312  ATCTCTCTGAGTTCAAGGCCATCTGAATTTACATAAAAGAGCTCCAGCCAGGGTG  3371

Qy  8542  cgcagtaaaacctgtctcaaaaaacaatctttagtgaccacgcaccc  8601
          |||||||||||||||||||||
Db  3372  CGCAGTAAAACCTTGTCTCAAAAAACAAAGCCATCTTTAGTGACCACAGGGCTTGCTCCACCCC  3431

Qy  8602  cagtgaccacggaccccacccccaccccgaccgtcgacggtcggggagacc  8661
          |||||||||||||||||||||
Db  3432  CAGTGACCACGGACCCCCACCCCACCCCGACCGCCGTGAGCCGCCCGGGGCCTGGAGGACC  3491

Qy  8662  agctgagtgtgcgctgggtcctcacacagcctcaagggatttcctcttccaagccaagt  8721
          |||||||||||||||||||||
Db  3492  AGCTGAGTGTGCGCTGGGGTCACACAGCTCTCAAGGATTTCCTCTTCCAAGCCAAGT  3551

Qy  8722  accagatccgctaccgtgggagggactgggagtgccgtccccgccccgga  8781
          |||||||||||||||||||||
Db  3552  ACCAGATCCGCTACCGGTGGGAGGGACACTTTGCCGCTCGGAAGGTCCCGCCGGA  3611

Qy  8782  ccgcccctgacccccgccccgatctgacctcctccacccgtgcaggtgga  8841
          |||||||||||||||||||||
Db  3612  CCAGTGCCTTGTGAACCCATTCGGGATCTCATTGGGTCGAAAAGGCCGGGAATCTGAGCCGA  3671

Qy  8842  cgtcagcaacaagaacctctcgcgtgagccgccgcacgtttacttcgt  8901
          |||||||||||||||||||||
Db  3672  CGTCAGCAACAGACCTCCTGCCCTCGGCCCTCTCGAAGCCAAGT  3731

Qy  8902  caaagtgcgttgtaaccattcgggattccggttcgaaaaagcgggaaatctgagcga  8961
          |||||||||||||||||||||
Db  3732  CCAAGTGCCTTGTGAACCCATTCGGGATCTCATTGGGTCGAAAAGGCCGGGAATCTGGACCGA  3791

Qy  8962  gtcggagcaacaaccgctgcctccaacccctccaggggctgg  9021
          |||||||||||||||||||||
Db  3792  GTGGGAGCGGCACCCCCACCGGTGCCTTCCCACCGGTGGTGAGCAACCTCTCCAGGGGCTGG  3851

Qy  9022  ctggcccaatgaatccaatccatcctgttccttcccccaccctttttgagacag  9081
          |||||||||||||||||||||
Db  3852  CTGGCCCAGGGAATCCCCAAATCCGTCCTTCCCCCACCGTTTTTGAGACAG  3911

Qy  9082  cgtcttcaaggtagcgcatgctggcccttaaattcaagtatgtagtcaaggatgacctcgagc  9141
          |||||||||||||||||||||
Db  3912  CGTCTTCAGGTAGGCATCGTGGCCTTAAATTCAGTAGTATCGTAAGGATGAACCTCGAGC  3971

Qy  9142  tcctgtcttttgtctccacttaggagacaagtggccatcacaccctttgggag  9201
          |||||||||||||||||||||
Db  3972  TCCTGGGTCTTTTGCTCCACTTAGAGACAAATGGCCCATCACCACCTTTGGGAG  4031

Qy  9202  actagcacatggagtctattaggtgacagatggagtacaacagtgtgac  9261
          |||||||||||||||||||||
Db  4032  ACTAGCCATGGAGTCTATTAGCCGTGTCATTTGGTGACAGATGGAGTACAACAGTGTGAC  4091

Qy  9262  ctcttcttaagagaacaagagctcgtttctcattcttaacccccaaatatccatggctctctagag  9321
          |||||||||||||||||||||
Db  4092  CTCTTGTTAAGAGAACTGGACAGGCTGTTTTTAACCCCAATAACCTAGGCTCTCTAGAG  4151

Qy  9322  gttaacttatatagaagaactgattatccacagaacct  9381
          |||||||||||||||||||||
Db  4152  GTTAACTTATATAAAAATAGAAGAACTATTATCACACAGAACCT  4211

Qy  9382  tttgtcacacacctatagaccacaatgccctgtcctaccacataagggtctcactagt  9441
          |||||||||||||||||||||
```

```
Db  4212  TTTGTCACACAACCTATAGACCACAGTGCCCTGTCCTACCCACACATAAGGGTCTTCTACTGCT  4271

Qy  9442  ggcccaccctccaacccctaaaagttaacctaggcagccttaataattgcaatcctcct  9501
          |||||||||||||||||||||
Db  4272  GGGCCCACCCTCCAACCCCCTAAAAAGGGTAACCTAGGCAGCCCTTAAATAATTGCAAATCCTCCT  4331

Qy  9502  acctcagaccttcttgatgctcaagaaaccaagcattaaaccaagttctctcttctcggggtc  9561
          |||||||||||||||||||||
Db  4332  ACCTCAGCCTTCTTGAAATGCTCAGAAACCAGCCATTAAACCCAAGTTTCTCTTCTCGGGTC  4391

Qy  9562  ccttcttaaggtgggagtgggggcctaaaagtgaccttcctttgtcctgaagagttctcatt  9621
          |||||||||||||||||||||
Db  4392  CCTTTCTTAAGGTGGGGAGTGGGGGCCTAAAGATGACTTCCTTTGTCCTGAAGAGTTCTCATT  4451

Qy  9622  catggatctgcactctctaatatgaacataactttagtgaccaggcttgctctctcaaatgaaaa  9681
          |||||||||||||||||||||
Db  4452  CATGGACATCGCACTCTCTAATATGAAATTGCATTAATAATTTCAGTGACTTCCGGAAAAATTA  4511

Qy  9682  cacctgtcaggtgtttaggcagcacagtcgggtccaagacacctcatttcggcaggcagtat  9741
          |||||||||||||||||||||
Db  4512  CACCTGGTTTAGGCAGCACAGTCGGGTCCAAGACACTTCATTATTTGCAGGCAGTAT  4571

Qy  9742  aagaaagaagtcccaatccccacacccaccccgcccccctctcccgggtccctaagaacagtcttaca  9801
          |||||||||||||||||||||
Db  4572  AAGAAGAAGTCCCATCCCCATCCCCACCGCTCCCCTAAGACCAGAGAATACTTCTACA  4631

Qy  9802  ctgaaactgaactctcgacgacgcatatgtctcaacttaatgatgataaaatggggaa  9861
          |||||||||||||||||||||
Db  4632  CTGAAACTGAACTCTCGCGACGCAATATGCCACTTTAAATGATGATGAAAATAAACTGGGGAA  4691

Qy  9862  acctgaggctccggagagagattcctggagaagagggtcaacaacctcattctgccaggcacttc  9921
          |||||||||||||||||||||
Db  4692  ACTGAGGCTCCGGAGAGATTCCTGGAGAAGAGGGTCAAAACCAGCTCCAATCTCC  4751

Qy  9922  agccccaatccgcgggctgttcctggggctcgggagttgacaacagttccca  9981
          |||||||||||||||||||||
Db  4752  AGCCCCAATCCGCGGGCTTCCTGTCCAGCCCGTCAGCGACTGGGAGG  4811

Qy  9982  gctggagccctgggagcttgtgctgtgccagcgctatatgtctccacttgcacggga  10041
          |||||||||||||||||||||
Db  4812  GCTGGAGCCCTGGGAGCTTTGGCCCTTGCCTGTCAGGGGCCTCGGGATTCTGCAGCACGGGGA  4871

Qy  10042 gccagacaggcggtcgcgtccagcccgcgccgagagaggaaccgggggttgggg  10101
          |||||||||||||||||||||
Db  4872  GCCAGCAGGGGACCCAAGGGCGGTCCCCAGAGAGGTCTAGGGTTGGGAGGGA  4931

Qy  10102 agtaggcaaggggcctgtgggggccggaaggcttgtctcaagcgagtcccagtt  10161
          |||||||||||||||||||||
Db  4932  GGTAAGCAGGGGCCTTGTGGGGGCCGAAGGCTTGTCTCAGCGAGTCCCCAGTT  4991

Qy  10162 ttattttatgccgtgaggcccgatgttcctatccgctgggggagtcggggct  10221
          |||||||||||||||||||||
Db  4992  TTATTTATGCCGTGAGGCCCGATGTCCTTATCCGCTGGGGGATGGCTCGGGCT  5051

Qy  10222 ggggattggaccccagggcttgggccttcccactcccactgtcaacacc  10281
          |||||||||||||||||||||
Db  5052  GGGGATTGGACCCCAAGGGCCTTCCCACTCCCACTGTCACACC  5111

Qy  10282 cgttgaattcctgaggctgaggggaaccgccctgttctctgctctgtctcttat  10341
          |||||||||||||||||||||
Db  5112  CGTTGCAATTCTCTGAGGCTTACCTTGGACCCGCCCTGTTCTGTCGTGTCTTCTAT  5171

Qy  10342 ttctgtcattcactttccagcgcctttttttttatgcttttttaatataataactacgttttaaa  10401
          |||||||||||||||||||||
Db  5172  TTCTGTCATTCACTTTCCAGCGCTTTTTTATGCTTTTTAATAATAAACTACGTTTTAAA  5231

Qy  10402 aattgcttttgtatatgtgtgctccacacaacacagtgaa  10461
          |||||||||||||||||||||
Db  5232  AATTGCTTTTGTATAATGTGTGCTCCACACACAACACAGTGAA  5291

Qy  10462 ggttagagaacttgttgatggctccctccaccatgtgggactaggggctgggcgacaag  10521
          |||||||||||||||||||||
Db  5292  GGTTAGAGAACTTGTTGTGACTAGGCTCCTTCCACCATGTGGGACTAGGCTGGGCCAACACAAG  5351
```

```
Qy  10522 agcaattactgagtcatctcgccacgcccctcacccctcactcttccatcctcgtttggatag 10581
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   5352 AGCAATTACTGAGTCATCTCGCCACGCCCCTCACCCCTCACTCCTCGTTTGGATAG 5411

Qy  10582 tcataggtaatcgaagtaaatcgctgggttctagctatcctggcctagcccta 10641
           |||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   5412 TCATAGGTAATCGAAGGTAAATCGCTGGGTTCTAGCTATCCTGCCTCAGCCTA 5471

Qy  10642 ccaagtgtcgtgtaccacagtttttgtggggagggggctcctcctccccagtgttgtgggt-aca 10700
           |||||||||||||||||||||||||||||||||||||||||||||| |||||| ||
Db   5472 CCAAGTGCTGTGTACCACACGTTTGTGGGGAGGGGGCTCTCCTCCCCAGTGTCTGGGGTGACA 5531

Qy  10701 cagtccccagaatcctgcttctctagtgtttcttgtcttagtttggcccttgtcttccgtg 10760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5532 CAGTCCCCAAGATCTCTGCTTTCTCAGTGTCTTTGTCTTAGTTTGGCCCTTGTCTTCCGTG 5591

Qy  10761 tccctagagtctccggccccaccttagtctccaattgagttcccttctgaccgaatactcgg 10820
           |||||| ||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db   5592 TCCCTAGAGTCTCCGGCCCCACCTTATCCAATTGACTAGGTCTTTCCTTACCGAATACTCGG 5651

Qy  10821 tttttacctcccactgacttgactccctccctccttgttgttccatcacgccctgtggcattgcca 10880
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   5652 TTTTTACCTCCCACTGATTTGACTCCCTCCTTTTGCTTGTCTCCATCGCCCGGTGGCANTTGCCA 5711

Qy  10881 ttcctctgggtgacttcctgggtccaacctgaacacttttcccaacctgcgaagc 10940
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5712 TTCCTCTCGGGTGACTTCTGGGTTCCAACCTGAACACTTTCCCAACCTGCGAAGC 5771

Qy  10941 tggctctggtatgggaggccgcccgtcctccgccagcccgcggcgccccaacact 11000
           |||  |||| |||||||||||||||||||||| ||||||||||||||||||
Db   5772 TGGTCTGGTATGGGAGGCCGCCCGTCCTGCGGCCCAACCTGGCCGCCCCAACACT 5831

Qy  11001 gcccgcccatctcttcttaaagcccgccgccggccggggggggggtgtgcagccgcgcggggc 11060
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   5832 GCCCGCTCCATTCCTCTTTAGACGCCCCGGGCCCGGCCGCGGGGGCCCGGGGTGCGAGCCGCGCGGGGC 5891

Qy  11061 ggcgagcccagctcgggcgtcgggccgcgagctcaagcagttcctcggctggctcaag 11120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5892 GGCGAGCCCAGCTCGGGCCGTCGGCGAGCTCAACAGTTCCTCGGCTGGCTCAAG 5951

Qy  11121 aagcacgcatactgctgtcgaaccttagttccgcctgtacgaccaagtggctcttggatg 11180
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   5952 AAGCACGCATACTGCCTCGAACCTTAGTTCGCGCTGTACGACCAGTGGCTGCTTGGATG 6011

Qy  11181 cagaagagtcacacagaccgaaacaagtaggaaagttggggagctgctgcgtggggggt 11240
           ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   6012 CAGAAGTCACACAAGACCCCGAAACCAGGTAGGAAAGTTGGGACGGCTTGCGTGGGGGT 6071

Qy  11241 aaaggagcaggagaggaagagagagagaacccgggtgagcagcccccacaacacctcttctt 11300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6072 AAAGGAGCAGGAGGAAGAGAGAGAGACCCGGGTGAGCAGCCTCCACAACACGCACTCTTCTT 6131

Qy  11301 tccaagcacacgaggacgagggggatcctcgccctcggggcgcagaggcgaaggtaag 11360
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   6132 TCCAAGCACACAGGACGACGAGGGGGATCCTCGCCCTGGGGCGCGCGGAGGGTAAG 6191

Qy  11361 gggtcgtggttgagtgggggcctacagcgtctagatgagggccccttctccctcctcggttgt 11420
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   6192 GGGTCTTGGGTGAGTGGGGCCTCACAGCGTCTAGATGAGGCCCCTTCCCCTCCTCGGTTGT 6251

Qy  11421 tgctcaaagggatctcttagtgtgtcattcaccaccactgcaaagagagccccagagttttacg 11480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6252 TGCTCAAAGGGGATCTTCATTTCACCACATGGCCTCAAAAGAAGCCCCAGGTTTTACTG 6311

Qy  11481 catcatcaagttgctgaaggatctgcggtccacgttacct 11540
           |||||||||||||||||||||||||||||||||||||||
Db   6312 CATCATCAAGTTTGTGAAAGGGTTCCAGGCTTAATGTGGCCCTCTCTTTTCTGCCCTCAGGTTTCCT 6371

Qy  11541 gccggctaaactctaaggataggccatcctctgcctgctgggtcaacctggaggctcactg 11600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6372 GCCGGCTAAACTCTAAGGATAGGCCATCCTCCTGCCTCGGCCTCAGACCTGGAGGCTCACCTG 6431
```

```
Qy  11601 aattggagccctctgtaccatctgggcaacaagaaaacctaccagagaggctggcacaat 11660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6432 AATTGGAGCCCTGTTACCATCTGGGCAACAACAACAAGGAAAACCTACCAGAGGCTGGGCACAAT 6491

Qy  11661 gagctcccacaacacagtcttggtcccacatgagtggtcacacaactggatataccccagtgt 11720
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   6492 GAGCTCCCACAACCACAGCTTTGGTCCCACATGAGTGGTCACACAACTGGATATACCCCAGTGT 6551

Qy  11721 gggtaggttggggtattgcaggggcctccacaagagtctcttcttaaatataaagaggttg 11780
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   6552 GGGTAGGGTTGGGGTATTGCAGGGGCCTCCAACAAGAGTCTCTTTAAATAAATAAAGAGTTG 6611

Qy  11781 ttcaggtcccgatggccagtgtgtttgtttggggcctatgtgctgggtgggggggga 11832
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6612 TTCAGGTCCCGATGGCCAGTGTGTTTGGGGGGCCTATGTGCTCTGGGGTGGGGGGA 6663
```

```
           Score 737.2; DB 1;  Length 1930;
           Pred. No. 1.7e-162;
   28;   Indels    3;  Gaps    2;

Qy  10983 tggccgcgccccaacactgccgctccattctcttagagcgccgggcccggggcgcggg 11042
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 TGGAGCCACCCCAACCACCCTGCCTCCACCCCTCGAAGTGAGCCCGGGGCCGCGGGG 1180

Qy  11043 gtgtgccaagccgcggggcgcgggagctgcgcgcgcggctgccaagcag 11102
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1181 GTGTGCGGAGCCGGGCGCGCGCGGAGCGGCCCCGAGCTCAAGCAG 1240

Qy  11103 ttcctcggctggctcaagaagcaccgcataactgctcgaacctagtgttccgcctgtacgac 11162
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1241 TTCCTCCGGCTGGCTCAAGAAGCACCGCATAACTGCTCGGAACCTTAGTTTCCGCCTGTACGAC 1300

Qy  11163 cagtggcgtgcttggatgcagaagtcacacaagaccgaaacaaggtaggaagttgggg 11222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1301 CAGTGGCGTGCTTGGATGCAGAAGTCACACAAGACCCGAAACCAGTAGTAGGAAAGTTGGGG 1360
```

```
CC   derivatives can be used for modulating the activity of the receptors e.g.
CC   to regulate development, maintenance or regeneration in an array of
CC   different cells and tissues in vitro and in vivo. They can be present in
CC   therapeutics used for modulating neuronal proliferation, differentiation
CC   and survival. The products can also be used for detection and diagnosis,
CC   e.g. for cancers or predisposition to cancers, or for drug screening.
SQ   Sequence 938 BP; 243 A; 245 C; 272 G; 178 T;
```

```
Query Match       6.28;  Score 737.2;  DB 1;  Length 938;
Best Local Similarity   96.2%;  Pred. No. 1.2e-162;
Matches 777;  Conservative 0;  Mismatches 28;  Indels 3;  Gaps 2;
```

```
Qy 10983 tggccgcgccgccccaaacactgccgctccattcctcttagagtccaagcccccgggcggg 11042
          |||  |||||  | || ||||||||||||||||||||||||||||| |||||||||||||||
Db   .79  TGGAGCGCCACCCCACCCGCTGCCCCTCCGAAGTGAGCCGCCCCGGGCCCCGGGCGGGGG 138

Qy 11043 gtgtgtgcaagccgccgcgggggcgggcgagcccaagctcggggccccgtgtgcgggcag 11102
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  139  GTGTGCGAGCGCCGCGGGGCGGGCCGAGCCCAGCTCGGGGCCCGTGCGGGGCTCAAGCAG 198

Qy 11103 ttcctcgctcggctcaagaagcacgcatactgtcgaacttagtttccgcctgtacgac 11162
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||| ||||
Db  199  TTCCTCGGCTCGGCTCAAGAAGCACACGCGGATACTGTCGAAACCTTAGTTTCCGCCTGTACGAC 258

Qy 11163 cagtggccgtgcttgatgcgaagtcacaacaagaccgaaacacaagtagagaaagttgggg 11222
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  259  CAGTGGCCGTGCTTGGATGCAGAAGTCACAAGACCCGAACCAGTGAGGAAAGTTGGGG 318

Qy 11223 gaggcttgcttgtgggggtaaaggagcagaggaagaggaagacccgggtgaagcagctcca 11282
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319  GAGGCTTGCTTGTGGGGGTAAAGGAGCAGAGGAAGAGGAAGACCCGGGTGAGCAGCTCCA 378

Qy 11283 caacaacgcactcctcttccaagcacacagagcagtctagataggccc 11342
          ||||||| | |||||||||||||||||| |||||||||||||||||||
Db  379  CAACACCGCACTCTTCTTCCAAGCACAGGAGGAATCCTGCCCTCGGGTCA 438

Qy 11343 gtgcggccgagaggtaagggggtctggggttgagttgtggggccctacagcagtctagataggagcccc 11402
          |||||||||||| ||||||||| ||||||||||||||||||||||||||||| ||||||||||
Db  439  GTGCGGCGGAGAGGTAAGGGGGCTGGGGCCCTACAGCAGTCTAGATGAGGCCC 498

Qy 11403 tttcccctccttcggtgttgctcaaaggatctcattcacccactgcaa 11462
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  499  TTTCCCCTCCTTCGGTGTTGCTCAAAGGGATCTCATTCACCCACTGCAAA 558

Qy 11463 gagcccaggtttactgcatcatcaagttgctgaaggttccaggcttaatgtggcctct 11522
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  559  GAGCCCCAGGTTTACTGCATCATCAAGTTGCTGAAGGGTCCAGGCTTAANTGTGGCCTCT 618

Qy 11523 ttttctgccctcaggtctcctgccggctaaactctaaggatagggccatctcctctgggtca 11582
          |r||||||||||||||||||||||||||||||| |||||||||| |||||||||||||||||
Db  619  TTTCTGCCCTCAGGTCTCCTGCCGGCTAAACTCTAAGGATAGGGCCATCTCCTCTGGGTCA 678

Qy 11583 gacctggaggctcacctgaattggagccccctctgtacc-atctgggcaacaaaagaaact 11641
          ||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db  679  GACCTGGAGGCTCACCTGAATTGGAGCCCCTCTGTACCTATCTGGGCAACAAAGAAACT 738

Qy 11642 acca--gaggctgggggcaaatgagctccacaaccacagctttggtccacatgatggtca 11699
          ||||  |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  739  ACCATGAGGCGTGGGGGCACAATGAGCTCCACAACCACAGCTTTGGTCCACATGATGGTCA 798

Qy 11700 cacttggatacccacgtgtgggtaggttggtagggcctcccaagagtctc 11759
          |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  799  CACTTGGATATACCCCAGTGTGGGTAGGTTGGGTATTGCAGGGCCTCCCAAGAGTCTC 858

Qy 11760 tttaaaataaataaagggggttgttcaggt 11787
          |||||||||||||||||||||||||||||||
Db  859  TTTAAATAAATAAAAGGAGTTGTTCAGGT 886
```

```
RESULT 5
V27140
ID  V27140 standard; cDNA; 1629 BP.
```

Second column (left of center):

```
CC   differentiation and survival of a wide variety of cells. The HR and it's
CC   The haemopoietin receptor (HR) NR6.3 is a form of the novel HR NR6.
CC   Interaction between the novel HR and a ligand facilitates proliferation,
PS   Claim 6; Page 90-92; 182pp; English.
PT   e.g. neuronal cells
PT   for modulating proliferation, differentiation and survival of cells,
PT   New isolated haemopoietin receptor - used for developing products
DR   P-PSDB; W55013.
DR   WPI; 98-260970/23.
PI   Zhang J;
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PA   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PA   (DZIE/) DZIEGLEWSKA H E.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PR   11-SEP-1996; AU-002246.
PD   19-MAR-1998.
PN   WO9811225-A2.
FH   Key            Location/Qualifiers
FT   CDS            1..468
FT                  /*tag= a
FT                  /product= "Haemopoietin receptor NR6.3"
OS   Mus sp.
OS   Mouse.
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ssi
KW   Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
DE   Novel haemopoietin receptor NR6.3 gene.
DT   29-SEP-1998 (first entry)
AC   V27142; AU-002246.
ID   V27142 standard; DNA; 938 BP.
V27142
RESULT 4
```

```
Db 1901 TTTAAATAAATAAAAGGAGTTGTTCAGGT 1928
Qy 1901 tttaaaataaataaaaggggttgttcaggt 11787
         |||||||||||||||||||||||||||||
Db 1841 CACTTGGATATACCCCAGTGTGGGTAAGGTTGGGTATTGCAGGGCCTCCCAACAAATCTC 1900
Qy 11700 cacttggatacccacgtgtgggtaggttggtagggcctcccaagagtctc 11759
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1781 ACCATGAGGCTGGGGCACAATGAGCTCCCACAACCACAGCTTTGGTCCACATGATGGTCA 1840
Qy 1781 acca--gaggctgggggcaaatgagctcccacaaccacagcttggtccacatgatggtca 11699
         ||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1721 GACCTGGAGGCTCACCTGAATTGGAGCCCCTCTGTACCTATCCCTATNTCGGGCAACAAAACCT 1780
Qy 1721 gacctggaggctcacctgaattggagcccctctgtacc-atctgggcaacaaaaggaaact 11641
         ||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db 1661 TTTCTGCCTCAGGTCTCCGGGCTAAACTCTAAGGATAGGGCCATCTCCTCTGGGTCA 1720
Qy 1661 ttttctgccctcaggtctcctgccggctaaactctaaggatagggccatctcctctgggtca 11582
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1601 GAGCCCCAGGTTTACTGCATCATCAAGTTGCTGAAGGGTCCAGGCTTAANTGTGGCCTCT 1660
Qy 1601 gagcccaggtttactgcatcatcaagttgctgaaggttccaggcttaatgtggcctct 11522
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1541 TTTCCCCTCCTTCGGTGTTGCTCAAAGGGATCTCATTCACCCACTGCAAA 1600
Qy 1541 tttcccctccttcggtgttgctcaaaggatctcattcacccactgcaa 11462
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1481 GTGCGGCGGAGAGGTAAGGGGGTCTGGGTGGGGCCTACAGCAGTCTAGATGAGGCCC 1540
Qy 11403 gtgcggccgagaggtaagggggtctggggttgggtggggccctacagcagtctagataggccc 11402
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1421 CAACACCCGCACTCTTCTTCCAAGCACAGAGCAGGGGATCCTGCCCTCGGGCAGACGGG 1480
Qy 11343 caacaacgcactcctcttccaagcacacagagcagtctagataggagcccc 11402
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1361 GAGGCTTGCTGGGGGGTCTGGGGGTAAGGGGGTCTGGGTGGGGCCTACAGCAGTCTAGATGAGGCC 1420
Qy 11223 gaggcttgcttgtgggggtaaaggagcagaggaagagggaagacccgggtgagcagctcca 11282
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1301 ACCATGAGGCTGGGGCACAATGAGCTCCCACAACCACAGCTTTGGTCCACATGATGGTCA 1840
Qy 11781 acca--gaggctgggggcaaatgagctcccacaaccacagcttggtccacatgatggtca 11699
         ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1721 CAGTGGCCGTGCTTGGATGCAGAAGTCACAAGACCCCTTGTACCTATCCCTATNTGGGCAACAAAACCT 1780
Qy 11163 cagtggccgtgcttgatgcgaagtcacaacaagaccgaaacacaagtagagaaagttgggg 11222
         ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 1661 TTCCTCGGCTCGGCTCAAGAAGCACACGCGGATACTGTCGAAACCTTAGTTTCCGCCTGTACGAC 1720
Qy 11103 ttcctcgctcggctcaagaagcacgcatactgtcgaacttagtttccgcctgtacgac 11162
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 1601 GTGTGCGAGCGCCGCGGGGCGGGCCGAGCCCAGCTCGGGGCCCGTGCGGGGCTCAAGCAG 198
Qy 11043 gtgtgtgcaagccgccgcgggggcgggcgagcccaagctcggggccccgtgtgcgggcag 11102
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1541 TGGAGCGCCACCCCACCCGCTGCCCCTCCGAAGTGAGCCGCCCCGGGCCCCGGGCGGGGG 138
Qy 10983 tggccgcgccgccccaaacactgccgctccattcctcttagagtccaagcccccgggcggg 11042
```

AC   V27140;
DT   29-SEP-1998 (first entry)
DE   Novel haemopoietin receptor NR6.1 gene.
KW   Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
OS   Mouse.
OS   Mus sp.
FH   Key   Location/Qualifiers
FT   CDS   113..1355
FT   /*tag= a
FT   /product= "Haemopoietin receptor NR6.1"
PN   WO9811225-A2.
PD   19-MAR-1998.
PF   11-SEP-1997; GO2479.
PR   11-SEP-1996; AU-002246.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PA   (DZIE/) DZIEGLEWSKA H E.
PI   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI   Zhang J;
DR   WPI; 98-260970/23.
DR   P-PSDB; W55011.
PI   New isolated haemopoietin receptor - used for developing products
PT   for modulating proliferation, differentiation and survival of cells,
PT   e.g. neuronal cells
PS   Claim 4; Page 77-81; 182pp; English.
CC   The haemopoietin receptor (HR) NR6.1 is a form of the novel HR NR6.
CC   Interaction between the novel HR and a ligand facilitates proliferation,
CC   differentiation and survival of a wide variety of cells. The HR and it's
CC   derivatives can be used for modulating the activity of the receptors e.g.
CC   to regulate development, maintenance or regeneration in an array of
CC   different cells and tissues in vitro and in vivo. They can be present in
CC   therapeutics used for modulating neuronal proliferation, differentiation
CC   and survival. The products can also be used for modulating neuronal
CC   e.g. for cancers or predisposition to cancers, or for drug screening.
SQ   Sequence 1629 BP; 336 A; 541 C; 453 G; 299 T;

Query Match   2.4%; Score 280.4; DB 1; Length 1629;
Best Local Similarity 99.6%; Pred. No. 6.9e-56;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   6330 acacagctgtaatacagccccccaggagaccaccacccttctcatcggctcccctgccaagcta 6389
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 ACACAGCTGTAATACAGCCCCCAGGAGACCACCACCCTTCTCATCGGCTCCCCTGCCAAGCTA 308

Qy   6390 ccctgctctatacatggagagacacacactgggggccaccgctgagggggctctactggacctca 6449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 CCTGCTCTATACATGGAGAGACACACACTGGGGGCCACCGCTGAGGGGGCTCTACTGGACCCTCA 367

Qy   6450 atggtcgccgcctgccctctgagctgtccgcgcctccttaacacctccaccctggccctgg 6509
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   368 ATGGTCGCCGCCTGCCCTCTGAGCTGTCCGCGCCTCCTTAACACCTCCACCCTGGCCCTGG 488

Qy   6510 ccctggctaaccttaatgggtccaggcaggcaagcaatctggtgtgtgtcacgccc 6569
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CCCTGGCTAACCTTAATGGGTCCAGGCAGGCAACAATCTGGTGTGTGTCACGCCC 487

Qy   6570 gagacggcagcattctggctggctcctgccctctatgttggct 6611
           |||||||||||||||||||||||||||||||||||||||||||
Db   488 GAGACGGCAGCAGCATTCTGGCTGGCTCCTGCCCTCTATGTTGGGCT 529

RESULT 6
ID   V27141 standard; cDNA; 1673 BP.
AC   V27141;
DT   29-SEP-1998 (first entry)
DE   Novel haemopoietin receptor NR6.2 gene.
KW   Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
OS   Mouse.
OS   Mus sp.

FH   Key   Location/Qualifiers
FT   CDS   1..1278
FT   /*tag= a
FT   /product= "Haemopoietin receptor NR6.2"
PN   WO9811225-A2.
PD   19-MAR-1998.
PF   11-SEP-1997; GO2479.
PR   11-SEP-1996; AU-002246.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PA   (DZIE/) DZIEGLEWSKA H E.
PI   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI   Zhang J;
DR   WPI; 98-260970/23.
DR   P-PSDB; W55012.
PI   New isolated haemopoietin receptor - used for developing products
PT   for modulating proliferation, differentiation and survival of cells,
PT   e.g. neuronal cells
PS   Claim 5; Page 84-87; 182pp; English.
CC   The haemopoietin receptor (HR) NR6.2 is a form of the novel HR NR6.
CC   Interaction between the novel HR and a ligand facilitates proliferation,
CC   differentiation and survival of a wide variety of cells. The HR and it's
CC   derivatives can be used for modulating the activity of the receptors e.g.
CC   to regulate development, maintenance or regeneration in vitro and in vivo.
CC   different cells and tissues in vitro and in vivo. They can be present in
CC   therapeutics used for modulating neuronal proliferation, differentiation
CC   and survival. The products can also be used for detection and diagnosis,
CC   e.g. for cancers or predisposition to cancers, or for drug screening.
SQ   Sequence 1673 BP; 344 A; 550 C; 474 G; 305 T;

Query Match   2.4%; Score 280.4; DB 1; Length 1673;
Best Local Similarity 99.6%; Pred. No. 7e-56;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   6330 acacagctgtaatacagccccccaggagaccaccacccttctcatcggctcccctgccaagcta 6389
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 ACACAGCTGTAATACAGCCCCCAGGAGACCACCACCCTTCTCATCGGCTCCCCTGCCAAGCTA 308

Qy   6390 ccctgctctatacatggagagacacacactgggggccaccgctgagggggctctactggacctca 6449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 CCTGCTCTATACATGGAGAGACACACACTGGGGGCCACCGCTGAGGGGGCTCTACTGGACCCTCA 368

Qy   6450 atggtcgccgcctgccctctgagctgtccgcgcctccttaacacctccaccctggccctgg 6509
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   369 ATGGTCGCCGCCTGCCCTCTGAGCTGTCCGCGCCTCCTTAACACCTCCACCCTGGCCCTGG 488

Qy   6510 ccctggctaaccttaatgggtccaggcaggcagcagtcaggaggacaatctggtgtgtgtcacgccc 6569
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   429 CCCTGGCTAACCTTAATGGGTCCAGGCAGGCAGCAGTCAGGAGGACAATCTGGTGTGTGTCACGCCC 488

Qy   6570 gagacggcagcattctggctggctcctgccctctatgttggct 6611
           |||||||||||||||||||||||||||||||||||||||||||
Db   489 GAGACGGCAGCAGCATTCTGGCTGGCTCCTGCCCTCTATGTTGGGCT 530

RESULT 7
ID   V41688 standard; cDNA; 1656 BP.
AC   V41688;
DT   26-OCT-1998 (first entry)
DE   Nucleotide sequence of the murine U4 gene.
KW   Murine; U4 protein; haematopoietin receptor superfamily;
KW   cell proliferation; immune response; antibody; cell differentiation;
OS   Mus sp.
KW   autoimmune disease; cancer; allergy; ds.
FH   Key   Location/Qualifiers
FT   CDS   122..1399
FT   /*tag= a
FT   /product= "U4 protein"
PN   WO9831811-A1.
PD   23-JUL-1998.
PF   15-JAN-1998; U00334.

PR   16-JAN-1997; US-784863.
PA   (GEMY) GENETICS INST INC.
PA   Collins M, Donaldson DD, Neben T, Whitters M;
PI   WPI; 98-414109/35.
DR   P-PSDB; W59804.
DR   New nucleic acid encoding U4 haematopoietin receptor superfamily
PT   chain - potentially useful, e.g. for modulating cell proliferation
PT   or immune response, for treating cancer and auto:immune disease
PS   Claim 1; Pages 25-26; 38pp; English.
CC   This is the nucleotide sequence encoding the murine U4 protein from
CC   the haematopoietin receptor superfamily, used in the method of the
CC   invention for the modulation of cell proliferation, or the immune
CC   response. Transformed mammalian cells are used to produce recombinant
CC   U4 protein. The U4 protein is used to screen for specific binding
CC   agents, raise antibodies. It is also used as reagents for assays and
CC   as tissue markers for isolation of cognate ligands and receptors, and
CC   in pharmaceutical compositions which may modulate cell proliferation,
CC   cell differentiation, and the immune system (e.g. for treating immune
CC   deficiency, inherited or the result of infection, autoimmune diseases,
CC   cancer, and allergy).
SQ   Sequence 1656 BP; 318 A; 552 C; 478 G; 308 T;

Query Match   2.3%; Score 277.2; DB 1; Length 1656;
Best Local Similarity 98.9%; Pred. No. 3.9e-55;
Matches 279; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6330 acacagcctgtaataacagcccccaggaagaaccccaccaccttctcatcggctctccctgcaagcta 6389
        |||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 246 ACACAGCTGTAATCAGCCCCCAGGAGACCCCACCTTTCTCATCGGCTCTCCCTGCAAGCTA 305

Qy 6390 cctgctctatacatggagacacacaccttgggggccacacgctgagggggtctcactggaccttca 6449
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 306 CCTGCTCTATACATGGAGACACACACCTTGGGGGCCACACGCTGAGGGGGTCTCACTGGACCCTCA 365

Qy 6450 atggtcgcgccgcccctcgagctgtcccgcccctcctaacaccctcacccctggcccctgg 6509
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 366 ATGGTCGCGCCGCCCCTCGAGCTGTCCCGCCCCTCCTTAACACCCTCACCCCTGGCCCCTGG 425

Qy 6510 ccctggctaaccttaaatgggtcagggcagcagtcaggagaacaatctggttgtgtcacgccc 6569
        ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db 426 CCCTGGCTAACCTTAATGGGTCAGGGCAGCAGTCAGGAGAACAATCTGGTTGTGTCACGCCC 485

Qy 6570 gagacggcagcagcattctggctggctcctcgcctctatcgttggct 6611
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 486 GAGACGGCAGCAGCATTCTGGCTGGCTCCTCGCCTCTCATAGTTGGCT 527

RESULT 8
ID   V27143 standard; cDNA; 834 BP.
AC   V27143;
DT   29-SEP-1998 (first entry)
DE   Nucleotide sequence of products generated by 5N race of brain cDNA.
KW   Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
OS   Mouse.
OS   Mus sp.
FH   Key   Location/Qualifiers
FT   CDS   1..834
FT   /*tag= a
FT   /product= "Haemopoietin receptor"
PN   WO9811225-A2.
PD   19-MAR-1998
PF   11-SEP-1997; G02479.
PR   11-SEP-1996; AU-002246.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PA   (DZIE) DZIEGLEWSKA H E.
PI   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI   Zhang J;
PI   WPI; 98-260970/23.
DR   P-PSDB; W55014.

PT   New isolated haemopoietin receptor - used for developing products
PT   for modulating proliferation, differentiation and survival of cells,
PT   e.g. neuronal cells
PS   Claim 7; Page 93-95; 182pp; English.
CC   The nucleotide sequence was generated by a 5N RACE of brain cDNA using
CC   NR6 specific primers. NR6 is a novel Haemopoietin receptor (HR).
CC   Interaction between the novel HR and a ligand facilitates proliferation,
CC   differentiation and survival of a wide variety of cells. The HR and it's
CC   derivatives can be used for modulating the activity of the receptors e.g.
CC   to regulate development, maintenance or regeneration in an array of
CC   different cells and tissues in vitro and in vivo. They can be present in
CC   therapeutics used for modulating neuronal proliferation, differentiation
CC   and survival. The products can also be used for detection and diagnosis,
CC   e.g. for cancers or predisposition to cancers, or for drug screening.
SQ   Sequence 834 BP; 167 A; 274 C; 225 G; 168 T;

Query Match   2.3%; Score 254.4; DB 1; Length 834;
Best Local Similarity 99.6%; Pred. No. 5.9e-50;
Matches 255; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 6356 cccacacccttctcatcggctctccctgcaagctacctgctctatacatggagacacacct 6415
        ||||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1 CCCACCCTTCTCATCGGCTCTCCCTGCAACTACCTGCTCTATACATGGAGACACACCT 60

Qy 6416 gggggccacacgctgaggggggtctcactggacccttcaatggtcgcgccgcccctcgagctg 6475
        |||| |||||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db   61 GGGGCCACACGCTGAGGGGGTCTCACTGGACCCTCAATGGTCGCGCCGCCCCTCGAGCTG 120

Qy 6476 tcccgcccctcctaacaccctcacccctggcccctggctaaccttaatggtccagg 6535
        |||||||||||| |||||||||||||||||||||||||||||||||||||||| ||
Db 121 TCCCGCCCCTCCTTAACACCCTCACCCCTGGCCCCTGGCTAACCTTAATGGGTCCAGG 180

Qy 6536 cagcagtcaggagaacaatctggttgtgtcacgcccgagacggcagcagcattctggctggctc 6595
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CAGCAGTCAGGAGAACAATCTGGTTGTGTCACGCCCGAGACGGCAGCAGCATTCTGGCTGGCTC 240

Qy 6596 tgcctctatgttggct 6611
        ||||||||||||||||
Db 241 TGCCTCTATGTTGGCT 256

RESULT 9
ID   V70896 standard; cDNA; 1724 BP.
AC   V70896;
DT   17-MAR-1999 (first entry)
DE   cDNA encoding rat Zcytor5.
KW   Zcytor5; cytokinin-like receptor; down-regulation; growth factor;
KW   maintenance factor; thyroid; heart; skeletal muscle; cardiotrophin-1;
KW   cardiac pathology; heart enlargement; Zcytor5 ligand; ss;
OS   Rattus sp.
FH   Key   Location/Qualifiers
     CDS   159..1436
FT   /*tag= a
FT   /product= Zcytor5
PN   WO9849307-A1.
PD   05-NOV-1998.
PF   01-MAY-1998; U08865.
PF   13-FEB-1998; US-074721.
PR   01-MAY-1997; US-045287.
PR   01-MAY-1997; US-850030.
PR   13-FEB-1998; US-023890.
PA   (ZYMO) ZYMOGENETICS INC.
PA   Adams RL, Foster DC, Gilbert T, Jelmberg AC, Lehner JM,
PI   Lok S, Presnell SR, Whitmore TE;
PI   WPI; 99-034662/03.
DR   P-PSDB; W70862.
DR   New mammalian cytokinin-like receptor Zcytor5 - useful for, e.g.
PT   down-regulating Zcytor5 natural ligands or detecting cardiotrophin-1
PT   in blood
PS   Disclosure; Page 72-75; 55pp; English.
CC   The present sequence encodes a protein designated Zcytor5, which is

CC   a cytokinin-like receptor. Soluble Zcytor5 may be administered to
CC   down-regulate the effects of a growth and/or maintenance factor in
CC   thyroid, heart, and skeletal muscle for example to lessen the effect
CC   of cardiotrophin-1 on cardiac pathologies, so preventing heart
CC   enlargement. Zcytor5 could be used to detect cardiotrophin-1 in the
CC   blood, and to discover other possible Zcytor5 ligands. A probe
CC   comprising Zcytor5 DNA or RNA can be used to determine the presence
CC   and integrity of the zcytor5 gene on chromosome 19. Antibodies and the
CC   anti-idiotypic antibody could be used to purify Zcytor5 and
CC   therapeutically to modify Zcytor5 ligand effects.
SQ   Sequence 1724 BP;   350 A;   550 C;   500 G;   324 T;

Query Match   2.1%;   Score 248;   DB 1;   Length 1724;
Best Local Similarity 92.9%;   Pred. No. 2.6e-48;
Matches 260;   Conservative 0;   Mismatches 20;   Indels 0;   Gaps 0;

Qy 6330 acacagctgtaatcagcccccaggagacccccacccttctcatcggctccctgcaagcta 6389
         ||||||||| |||||| ||||||||||||||| |||||||||||||| ||||||| |||||
Db   283 ACACAGCTGTGATCAGTCCCCAGGAGACCCCACTCTTCTCATCGGATCCTCCCTTCARGCTA 342

Qy 6390 cctgctcctatacatggagaacacacctgggggccacccgctgagggggctctactggaccttca 6449
         ||||||||| | |||||||||||||||||||||||||||||||||||||||||| | ||| ||||
Db   343 CGTGCTCTATACATGGAGAACACACCGGGGGCCACACGCTGAGGGGCTTCTACTGGACCCTCA 402

Qy 6450 atggtcgtccgccgcctctgagctgtcccgccctcttaacacctccacccctggccctgg 6509
         | ||||| |||| |||||| |||||| |||||||| |||||||||||||||||||||||||
Db   403 ACGGCGGCCGCCTGCCCTCAGAGCTGTCCCGTCTCCTCAACACCTCCACCCTGGCCTTGG 462

Qy 6510 ccctggctaaccttaatgggtccaggcagcagtcaggaggagacaatctggttgtgtcacgccc 6569
         ||||||| |||||||||||| ||||| ||||| |||||||||||||||||||||||||||||
Db   463 CCCTGGCTAACCTTAATGGGTCCAGGCAGCAGTCAGGGGACAAATCTGGTTGTGTCACGCCC 522

Qy 6570 gagacggcagcgaattctggctggctcctgccctctatgttgg 6609
         |||||||||||||||||||||||||||||||||||||||||||
Db   523 GAGATGGGCACGCAATTCTGGCTGGTTCCTGCCTCTATGTTGG 562

PS   Disclosure: Page 63-66; 55pp; English.
CC   The present sequence encodes a protein designated zcytor5, which is
CC   a cytokinin-like receptor. Soluble Zcytor5 may be administered to
CC   down-regulate the effects of a growth and/or maintenance factor in
CC   thyroid, heart, and skeletal muscle for example to lessen the effect
CC   of cardiotrophin-1 on cardiac pathologies, so preventing heart

---

RESULT 11

PS   Disclosure: Page 68-70; 55pp; English.
CC   The present sequence encodes an allelic variant of protein designated
CC   zcytor5, which is a cytokinin-like receptor. Soluble Zcytor5 may be
CC   administered to down-regulate the effects of a growth and/or maintenance
CC   factor in thyroid, heart, and skeletal muscle for example to lessen the
CC   effect of cardiotrophin-1 on cardiac pathologies, so preventing heart
CC   enlargement. Zcytor5 could be used to detect cardiotrophin-1 in the
CC   blood, and to discover other possible Zcytor5 ligands. A probe
CC   comprising Zcytor5 DNA or RNA can be used to determine the presence

---

Query Match   1.9%;   Score 221.8;   DB 1;   Length 1690;
Best Local Similarity 86.8%;   Pred. No. 3.4e-42;
Matches 244;   Conservative 0;   Mismatches 37;   Indels 0;   Gaps 0;

Qy 6330 acacagctgtaatcagcccccaggagacccccacccttctcatcggctccctgcaagcta 6389
         ||||||||| |||| |||||||||||||||||| ||||||| |||||| ||||| |||| |
Db   167 ACACAGCTGTGATCAGTGAATCAGTCCCCAGGAGATCCCACAGCCTTCTCATCGGCCTTCCTTGGCCA 226

Qy 6390 atggtcgtccgccgcctctgagctgtcccgccctcttaacacctccacccctggccctgg 6449
         | ||||| |||| |||||| |||||| || |||||||| || |||||||||||||||||
Db   227 CCTGCTCAGTCGCACGGAGACCCCACCAGGAGACCACCACGGGCCTCTACTGGACCCTCA 286

Qy 6450 atggtcgtccgccgcctctgagctgtcccgccctcttaacacctccacccctggccctgg 6509
         |||||| |||| |||||| |||||| || || ||||||| || ||| |||||||||||||
Db   287 ATGGGCGCCGCGCCCCGCTCCCCTGAGCTCTCCCGTACTCACAAGCCCGTACTGGCCTCTGG 346

Qy 6510 ccctggctaaccttaatgggtccaggcagcagtcaggaggagacaatctggttgtgtcacgccc 6569
         |||||||| ||| ||||| ||||| |||||| |||||||||||||||||||||||||||||
Db   347 CCCTGGCCAACCTCAATGGGTTCAGGCAGCAGTCAGGGGACAAACTCGGTTGTGTCCCAGCGCC 406

Qy 6570 gagacggcagcgaattctggctggctcctgccctctatgttggc 6610
         |||||||||||||||||||||||||||||||||||||||||||
Db   407 GTGACGGCAGCATTCTGGCTGGCTCCTGCCTCTCTATGTTGGC 447

```
CC  and integrity of the zcytor5 gene on chromosome 19. Antibodies and the
CC  anti-idiotypic antibody could be used to purify zcytor5 and
CC  therapeutically to modify zcytor5 ligand effects.
SQ  Sequence  1813 BP;  415 A;  604 C;  519 G;  275 T;

Query Match    1.9%;  Score 221.8; DB 1;  Length 1813;
Best Local Similarity  86.8%;  Pred. No. 3.5e-42;
Matches  244;  Conservative  0;  Mismatches  37;  Indels  0;  Gaps  0;

Qy  6330 acacagctgtaatcagccccaggagacccaccctttctcatcggctctccctgcaagcta 6389
         |||||||| ||||||| |||||||||||||||||| |||||||||||||||||| ||| |
Db   200 ACACAGCTGTGATCAGTCCCCAGGAATCCCACGCTTCTCATCGGCTCTCCCTGCGGCCA 259

Qy  6390 cctgtctctacatggagacacacctggggcgccacgcgctgagggggtctcactggaccttca 6449
         |||||| || |||||||||||||||||| |||||||| |||||||||||||||||||| ||
Db   260 CCTGCTCAGTGACGGAGGACCCACCAGGAGCCACCGCCGGGGGCCTTCTACTGGACCCTCA 319

Qy  6450 atggtccgccgcctgagctcgtgtccccgcctccttaacacctcaccacccttggcccctgg 6509
         |||| |||||||||||||||||||||||||||| ||||||||||||||||||| ||||| |
Db   320 ATGGGCCGCCGGCGTCGCCCGGCCTCTCCCGCGTACTCAACACGCCTCCACCTTGGCCTCGG 379

Qy  6510 ccctggctaaccttaatgggtcaggcagtcagtgaacaatctggtgtgtcacgccc 6569
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   380 CCCTGGCCAACCTCAATGGGTCCAGGCAGGCACAACCTCGTGTGTGCCACGCCC 439

Qy  6570 gagacggcagcagcattctcgtctggctctcgccctctatgttggc 6610
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   440 GTGACGGCGAGCARCCTGGCTCGTCGGTCGGCTCCTCATCTATGTTGGC 480

RESULT 12
ID  V41689
AC  V41689;
DT  26-OCT-1998  (first entry)
DE  Human: U4 protein: haematopoietin receptor superfamily;
DE  Nucleotide sequence of the human U4 gene.
KW  cell proliferation; immune response; antibody; cell differentiation;
KW  autoimmune disease; cancer; allergy; ds.
OS  Homo sapiens.
FH  Key             Location/Qualifiers
FT  CDS             1..1228
FT                  /*tag= a
FT                  /product= "U4 protein"
PN  WO9831811-A1.
PD  23-JUL-1998.
PF  15-JAN-1998; U00334.
PR  16-JAN-1997; US-784863.
PA  (GEMY ) GENETICS INST INC.
PI  Collins M, Donaldson DD, Neben T, Whitters M;
DR  WPI: 98-414109/35.
DR  P-PSDB: W59805.
PT  New nucleic acid encoding U4 haematopoietin receptor superfamily
PT  chain - potentially useful, e.g. for modulating cell proliferation
PT  or immune response, for treating cancer and auto:immune disease
PS  Claim 1; Page 28; 38pp; English.
CC  This is the nucleotide sequence encoding the human U4 protein from
CC  the haematopoietin receptor superfamily, used in the method of the
CC  invention for the modulation of cell proliferation, or the immune
CC  response. Transformed mammalian cells are used to produce recombinant
CC  U4 protein. The U4 protein is used to screen for specific binding
CC  agents, raise antibodies. It is also used as reagents for assays and
CC  as tissue markers for isolation of cognate ligands and receptors, and
CC  in pharmaceutical compositions which may modulate cell proliferation,
CC  cell differentiation, and the immune system (e.g. for treating immune
CC  deficiency, inherited or the result of infection, autoimmune diseases,
CC  cancer, and allergy).
SQ  Sequence  1579 BP;  304 A;  535 C;  473 G;  267 T;

Query Match    1.9%;  Score 220.2; DB 1;  Length 1579;
Best Local Similarity  86.5%;  Pred. No. 7.7e-42;
```

```
Matches  243;  Conservative  0;  Mismatches  38;  Indels  0;  Gaps  0;

Qy  6330 acacagctgtaatcagccccaggagacccaccctttctcatcggctctccctgcaagcta 6389
         |||||| |||||||||||||||||||||| |||||||||||||||||||| ||| |
Db    75 ACACAGCTGTGATCAGTCCCCAGGAATCCCACGCTTCTCATCGGCTCTCCCTGCGGCCA 134

Qy  6390 cctgtctctacatggagacacacctggggcgccacgcgctgagggggtctcactggaccttca 6449
         |||||| || ||||||||||||||||| |||||||| |||||||||||||||||||| ||
Db   135 CCTGCTCAGTGACGGAGGACCCACCAGGAGCCACCGCCGGGGGCCTTCTACTGGACCCTCA 194

Qy  6450 atggtccgccgcctgagctcgtgtccccgcctccttaacacctcaccacccttggcccctgg 6509
         |||| ||||||||||| ||||||||||||||| |||||||||||||||||||| ||||| |
Db   195 ACGGGCCGCCGGCGTCGCCCGGCCTCTCCCGAGCTCTCAACACGCCTCCACCTTGGCCTCGG 254

Qy  6510 ccctggctaaccttaatgggtcaggcagtcagtgaacaatctggtgtgtcacgccc 6569
         |||||| ||||||||||||||||||||||||| |||||||||||||||||||||
Db   255 CCCTGGCCAACCTCAATGGGTCCAGGCAGGCACAACCTCGTGTGTGCCACGCCC 314

Qy  6570 gagacggcagcagcattctcgtctggctctcgccctctatgttggc 6610
         ||||||||||| |||||||||||||||||||||||||||||||||||
Db   315 GTGACGGCGAGCAATCCTGGCTCGGTCGGTCCTCTCTACGTTGGC 355

RESULT 13
ID  V27144
AC  V27144;
DT  29-SEP-1998  (first entry)
DE  Nucleotide sequence of clone HFK-66 encoding human NR6.
KW  Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW  cell survival; therapeutic; neuronal proliferation; drug screening; ss;
KW  Human.
OS  Homo sapiens.
FH  Key             Location/Qualifiers
FT  CDS             1..1053
FT                  /*tag= a
FT                  /product= "Human NR6"
PN  WO9811225-A2.
PD  19-MAR-1998.
PF  11-SEP-1997; G02479.
PR  11-SEP-1996; AU-002246.
PA  (AMRA-) AMRAD OPERATIONS PTY LTD.
PA  (DZIE/) DZIEGIELEWSKA H E.
PI  Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI  Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI  Zhang J;
DR  WPI: 98-260970/23.
DR  P-PSDB: W55015.
PT  New isolated haemopoietin receptor - used for developing products
PT  for modulating proliferation, differentiation and survival of cells,
PT  e.g. neuronal cells
PS  Claim 7; Page 102-104; 182pp; English.
CC  The NR6 gene encodes a novel Haemopoietin receptor (HR). Interaction
CC  between the novel HR and a ligand facilitates proliferation,
CC  differentiation and survival of a wide variety of cells. The HR and it's
CC  derivatives can be used for modulating the activity of the receptors e.g.
CC  to regulate development, maintenance or regeneration in an array of
CC  different cells and tissues in vitro and in vivo. They can be present in
CC  therapeutics used for modulating neuronal proliferation, differentiation
CC  and survival. The products can also be used for detection and diagnosis,
CC  e.g. for cancers or predisposition to cancers, or for drug screening.
SQ  Sequence  1391 BP;  281 A;  459 C;  417 G;  234 T;

Query Match    1.8%;  Score 190.8; DB 1;  Length 1391;
Best Local Similarity  73.5%;  Pred. No. 5.3e-35;
Matches  291;  Conservative  0;  Mismatches  37;  Indels  68;  Gaps  1;

Qy  8604 gtgaccacggacaccccaccgacgtgcacgtgagcccgttcggggggcctggggggaggacag 8663
         |||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db   469 GTGACCACGGACCCCCCGACGTGGAGCCCCGTCGTGGGGCCTGGGGGAGGACCAG 528

Qy  8664 ctgagtgtgcgcgtgggctccaccagcctcaaggattcctcttccaagccaagtac 8723
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   529 ctgagtgtgcgcgtgggctccaccagcctcaaggattcctcttccaagccaagtac ...
```

Db 529 CTGAGCGTCGCGCTGGGTGTCGCGCCGCCTCAAGGATTCCTTCTTTCAAGCCAAATAC 588
         |||| |||||||| |||||||||| || ||||||||||| ||||||||||||||| |||

Qy 589 CAGATCCGCTACCGAGTGGAGGACAGTGTGGGATGGAAGGTG 626

Qy 8724 cagatccgctaccgcgctgggtggaggacagcgtggagaggtgcggggtcccgccccggaccaccccgccctcaaggattcccttcttcaagccaaatac 8783
          |||||||||||| ||| |||| || ||||||| |||||||||||| |||||||||||||||| ||||||||||||||||||||||||| ||||||||||||

Db 589 CAGATCCGCTACCGAGTGGAGGACAGTGTGGGATGGA 626

Qy 8784 cgccctgaccccgcccccccgcatctgactccctcaccgtgcaggtggtggatgacg 8843
          --AGGTGGTGGACGATG 640

Db 626 --AGGTGGTGGACGATG 640

Qy 8844 tcagcaaccagacctctctgccgtctcgcgggcctgagcgcacgtttacttcgtcc 8903
          |||||||| || |||||||| | |||| |||||||| ||||||| ||||||||||||

Db 641 TGACAACCAGACCTCTCCGCCGCCTGGCCACCGTGTACTTCGTGC 700

Qy 8904 aagtgcgttgtaaccattccggatctattgggtcgaaaaggcgggaatctggagcgagt 8963
          |||||||||| ||||||| |||||| ||||||| |||||| ||||| |||||| | ||||

Db 701 AAGTGCGTCGCCAACCCTTTGGCATCTRATGGCTCCAAGAAAGCCGGGATCTGGAGTGAGT 760

Qy 8964 ggagccacccaccgctgcctccacccctcgaagtg 8999
          |||||||| |||||| ||| ||||| || ||||||

Db 761 GGAGCCACCCACACGGCGCCCTTCACCTCCCGCAGTG 796

RESULT 14
ID   V27159
V27159; standard; DNA; 560 BP.
AC   V27159;
DT   29-SEP-1998 (first entry)
DE   PCR product for human NR6.
KW   Haemopoietin receptor; cell proliferation; cell differentiation; cancer;
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
KW   Human.
OS   Homo sapiens.
PN   WO9811225-A2.
PD   19-MAR-1998.
PF   11-SEP-1997; G02479.
PR   11-SEP-1996; AU-002246.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PA   (DZIE/) DZIEGLEWSKA H E.
PI   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI   Zhang J;
DR   WPI; 98-260970/23.
PT   New isolated haemopoietin receptor - used for developing products
PT   for modulating proliferation, differentiation and survival of cells,
PT   e.g. neuronal cells.
PS   Disclosure; Page 101; 182pp; English.
CC   NR6 is a novel haemopoietin receptor (HR). Interaction between the novel
CC   HR and a ligand facilitates proliferation, differentiation and survival
CC   of a wide variety of cells. The HR and it's derivatives can be used for
CC   modulating the activity of the receptors e.g. to regulate development,
CC   maintenance or regeneration in an array of different cells and tissues in
CC   vitro and in vivo. They can be present in therapeutics used for
CC   modulating neuronal proliferation, differentiation and survival. The
CC   products can also be used for detection and diagnosis, e.g. for cancers
CC   or predisposition to cancers, or for drug screening.
SQ   Sequence 560 BP;

Query Match   1.2%; Score 141.2; DB 1; Length 560;
Best Local Similarity   92.0%;
Matches 149; Conservative   0; Mismatches 13; Indels   0; Gaps   0;

Qy 8604 gtgaccaacggaccccccaccccgacgtgcacgtgagccgcgtgggggcctgagggaccag 8663
          ||||| |||||| || |||| ||||| |||| ||||||||||| ||||| |||||||||||

Db 382 GTGACCACGGACCCCCCGCCGACGTGCACGTGAGCCGCGTGGGGGCCTGAGGGACCAG 441

Qy 8664 ctgagtgtgtcgcgtggtctccacacagctccaaggattcctcttccaagccaagtac 8723
          |||| | ||||| ||| |||||| ||| || |||||||||||||| |||||||| |||

Db 442 CTGAGCGTGCGCTGGGGTGTCGCCACACCCCCTCAAGGATTTCCTTTTTTCAAGCCAAATAC 501

Qy 8724 cagatccgctaccgcgtgggagacagcgtggactggaggagaagcggggtcccgccccggaccaccccgccctcaaggattcccttcttcaagccaaatac 8823

RESULT 15
ID   V27157
V27157; standard; DNA; 143 BP.
AC   V27157;
DT   29-SEP-1998 (first entry)
DE   Nucleotide sequence unique to 5N RACE of brain cDNA.
KW   Haemopoietin receptor; cell proliferation; cell differentiation;
KW   cell survival; therapeutic; neuronal proliferation; drug screening; ss;
OS   Mus sp.
OS   Mouse.
Key   Location/Qualifiers
CDS   105..143
           /tag= a
           /note= "No stop codon given"
PN   WO9811225-A2.
FT   19-MAR-1998.
FT   11-SEP-1997; G02479.
FT   11-SEP-1996; AU-002246.
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
PA   (DZIE/) DZIEGLEWSKA H E.
PI   Alexander W, Fabri L, Farley A, Hilton DJ, Kikuchi Y,
PI   Kojima T, Maeda M, Nash A, Nicola NA, Rakar S, Willson T,
PI   Zhang J;
DR   WPI; 98-260970/23.
DR   P-PSDB; W55018.
PT   New isolated haemopoietin receptor - used for developing products
PT   for modulating proliferation, differentiation and survival of cells,
PT   e.g. neuronal cells.
PS   Disclosure; Page 98; 182pp; English.
CC   The nucleotide sequence is unique to brain V27157. NR6 is a novel
CC   haemopoietin receptor (HR). Interaction between the novel HR and a
CC   ligand facilitates proliferation, differentiation and survival of a wide
CC   variety of cells. The HR and it's derivatives can be used for modulating
CC   the activity of the receptors e.g. to regulate development, maintenance
CC   or regeneration in an array of different cells and tissues in vitro and
CC   in vivo. They can be present in therapeutics used for modulating
CC   neuronal proliferation, differentiation and survival. The products can
CC   also be used for detection and diagnosis, e.g. for cancers or
CC   predisposition to cancers, or for drug screening.
SQ   Sequence 143 BP;

Db 502 CAGATCCGCTACCGAGTGGAGGACAGTGTGGGATGGAAGGTG 543
         ||||||||||||| |||||||||||||| |||||| |||||||

Query Match   1.0%; Score 121.2; DB 1; Length 143;
Best Local Similarity   97.6%;
Matches 123; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy 331 ggcatgaaggcttagggtggggatcggtaggacccatgcacccagagaaggaaggggactggtg 390
         ||||||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||

Db 1   GGCATGAAGGCTTAGGGTGGGGATCGGTAGGACCCATGCACCCAGAGAAAGGAAGGGGACTGGTG 60

Qy 391 gcaacttcaaatctctgtgggaaggaagagggctgaagaggaatgaacggtcagt 450
         ||||||| |||| ||||| ||||||||||||||| ||||||||||||| ||||||

Db 61   GCAACTTTCAAACTCTCTGGGGAAGAAGGGCTGAAGAGGAATGAACGGGCTCAGAC 120

Qy 451 actgct 456
         || |||

Db 121 ACAGCT 126

Search completed: September 19, 1999, 16:19:49
Job time: 16783 sec

~ NAME: Misrock, S Leslie
~ REGISTRATION NUMBER: 18,872
~ REFERENCE/DOCKET NUMBER: 7005-030
~ TELECOMMUNICATION INFORMATION:
~ TELEPHONE: 212 790-9090
~ TELEFAX: 212 869864/9741
~ TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 1355 base pairs
~ TYPE: nucleic acid
~ STRANDEDNESS: single
~ TOPOLOGY: linear
~ MOLECULE TYPE: cDNA
US-08-442-281-4

Query Match
Best Local Similarity 78.4%;   Score 99; DB 3; Length 1355;
Matches 131; Conservative 0; Pred. No. 6.7e-15; Mismatches 35; Indels 1; Gaps 1;

Qy 5533 gtttgtattggttggttggttttggagaacaaggcttctctgtgt 5592
Db 1355 GTGTTCATTTTGTTGTTGTTTGTTGTTTTTCGAGACAGGGTTTCTGTGT 1296

Qy 5593 agtcctgggtgtccttggaactcactctgtagacacaggcgtggccttgaactcagaaatcc 5652
Db 1295 AGCTCTGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttgtgtgtctcccaagtgtcttagaattaaggtgtgcactgcca 5699
Db 1236 TCCGTCTGTGCATCCCAGGTGCTGGGTATTAAAGGCGAGCTCTACCA 1190

RESULT 11
US-08-442-281-16/c
Sequence 16, Application US/08442281
Patent No. 5807991
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,281
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/757,390
FILING DATE: 10-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:

~ SEQUENCE CHARACTERISTICS:
~ LENGTH: 1355 base pairs
~ TYPE: nucleic acid
~ STRANDEDNESS: single
~ TOPOLOGY: linear
~ MOLECULE TYPE: cDNA
~ FEATURE:
~ NAME/KEY: CDS
~ LOCATION: 13..1008
US-08-442-281-16

Query Match
Best Local Similarity 78.4%;   Score 99; DB 3; Length 1355;
Matches 131; Conservative 0; Pred. No. 6.7e-15; Mismatches 35; Indels 1; Gaps 1;

Qy 5533 gtttgtattggttggttggttttggagaacaaggcttctctgtgt 5592
Db 1355 GTGTTCATTTTGTTGTTGTTTGTTGTTTTTCGAGACAGGGTTTCTGTGT 1296

Qy 5593 agtcctgggtgtccttggaactcactctgtagacacaggcgtggccttgaactcagaaatcc 5652
Db 1295 AGCTCTGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttgtgtgtctcccaagtgtcttagaattaaggtgtgcactgcca 5699
Db 1236 TCCGTCTGTGCATCCCAGGTGCTGGGTATTAAAGGCGAGCTCTACCA 1190

RESULT 12
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.

WLPOX VIRUS

e 500

...Version #1.25

APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENI, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/442,282
  FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/757,390
  FILING DATE: 10-SEP-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Misrock, S. Leslie
  REGISTRATION NUMBER: 18,872
  REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 212 790-9090
  TELEFAX: 212 8698864/9741
  TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1355 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-442-282-4

Query Match
Best Local Similarity   78.4%;
Matches 131;   Conservative 0;   Mismatches 35;   Indels 1;   Gaps 1;

Qy 5533 gtttgtatttggttggttcgttgagtttgttttattggagacaggggcttctctgtgt 5592
         ||||| ||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 1355 GTGTTCATTTTTGTTTGTTTTGTTTTTCGAGACAGGGTTCTCTGTGT 1296

Qy 5593 agtcctggctgtccttggaactcactctgtagacccttgaactcagaaatcc 5652
         || ||||||||||||||||||||||||||| || |||||| |||||||||||
Db 1295 AGCTCTGGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttgtgctccccaagtgcttagattaaaggtgtgcaactgcca 5699
         ||||| ||||| ||||||| |||||||||||| ||||||||||||||||
Db 1236 TCCCGTCTGTGCAATCCCAGGTGCTGGTATTAAAGGCGAGCTCTACCA 1190

RESULT 9
US-08-442-282-16/c
Sequence 16, Application US/08442282
Patent No. 5760204
GENERAL INFORMATION:
  APPLICANT: Takatsu, Kiyoshi
  APPLICANT: Tominaga, Akira
  APPLICANT: Takagi, Satoshi
  APPLICANT: Murata, Yoshiyuki
  TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
  NUMBER OF SEQUENCES: 18
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Pennie & Edmonds
    STREET: 1155 Avenue of the Americas
    CITY: New York
    STATE: New York
    COUNTRY: U.S.A.
    ZIP: 10036-2711
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/442,282
    FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/757,390
  FILING DATE: 10-SEP-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Misrock, S. Leslie
  REGISTRATION NUMBER: 18,872
  REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 212 790-9090
  TELEFAX: 212 8698864/9741
  TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1355 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
  NAME/KEY: CDS
  LOCATION: 13..1008
US-08-442-282-16

Query Match
Best Local Similarity   78.4%;
Matches 131;   Conservative 0;   Mismatches 35;   Indels 1;   Gaps 1;

Qy 5533 gtttgtatttggttggttcgttgagtttgttttattggagacaggggcttctctgtgt 5592
         ||||| ||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 1355 GTGTTCATTTTTGTTTGTTTTGTTTTTCGAGACAGGGTTCTCTGTGT 1296

Qy 5593 agtcctggctgtccttggaactcactctgtagacccttgaactcagaaatcc 5652
         || ||||||||||||||||||||||||||| || |||||| |||||||||||
Db 1295 AGCTCTGGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttgtgctccccaagtgcttagattaaaggtgtgcaactgcca 5699
         ||||| ||||| ||||||| |||||||||||| ||||||||||||||||
Db 1236 TCCCGTCTGTGCAATCCCAGGTGCTGGTATTAAAGGCGAGCTCTACCA 1190

RESULT 10
US-08-442-281-4/c
Sequence 4, Application US/08442281
Patent No. 5807991
GENERAL INFORMATION:
  APPLICANT: Takatsu, Kiyoshi
  APPLICANT: Tominaga, Akira
  APPLICANT: Takagi, Satoshi
  APPLICANT: Murata, Yoshiyuki
  TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
  NUMBER OF SEQUENCES: 18
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Pennie & Edmonds
    STREET: 1155 Avenue of the Americas
    CITY: New York
    STATE: New York
    COUNTRY: U.S.A.
    ZIP: 10036-2711
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/442,281
    FILING DATE:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/757,390
    FILING DATE: 10-SEP-1991
ATTORNEY/AGENT INFORMATION:

RESULT 6
US-07-757-390-4/c
Sequence 4, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1355 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-757-390-4

Query Match   0.8%;   Score 99;   DB 1;   Length 1355;
Best Local Similarity   78.4%;   Pred. No. 6.7e-15;
Matches 131; Conservative   0;   Mismatches 35;   Indels   1;   Gaps   1;

Qy 5533 gtttgtattggttcggttggagtttgttattggagacaagggcttcctgtgt 5592
        |||| ||||| ||||||||||||| || ||||||||||||| |||| |||||
Db 1355 GTGTTCATTTTGTTTTGTTTGTTTGTTTTTCGAGACAGGGTTTCTCTGTGT 1296

Qy 5593 agtcctgcttgtcctggaactcactctgtagacaccagggcctgacctgga 5652
        || ||||| ||||| |||||||||| | ||||||| || ||||| |||||||
Db 1295 AGCTCTGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttggtgcttcccaagtgcttagattaaaggtgtgcactgcca 5699
        ||| ||| || |||||||||||| ||||||||||||| ||||||| ||
Db 1236 TCCCGTCTGTGCATCCCAGGTGCTGGTATTAAAGGCGAGCTCACCA 1190

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1355 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1008
US-07-757-390-16

Query Match   0.8%;   Score 99;   DB 1;   Length 1355;
Best Local Similarity   78.4%;   Pred. No. 6.7e-15;
Matches 131; Conservative   0;   Mismatches 35;   Indels   1;   Gaps   1;

Qy 5533 gtttgtattggttcggttggagtttgttattggagacaagggcttcctgtgt 5592
        |||| ||||| ||||||||||||| || ||||||||||||| |||| |||||
Db 1355 GTGTTCATTTTGTTTTGTTTGTTTGTTTTTCGAGACAGGGTTTCTCTGTGT 1296

Qy 5593 agtcctgcttgtcctggaactcactctgtagacaccagggcctgacctgga 5652
        || ||||| ||||| |||||||||| | ||||||| || ||||| |||||||
Db 1295 AGCTCTGGCTGTCC-TGGAACTCACTTTATAGACCAAGCTGGCCTCGAACTCAGAAATCC 1237

Qy 5653 gcctgcttggtgcttcccaagtgcttagattaaaggtgtgcactgcca 5699
        ||| ||| || |||||||||||| ||||||||||||| ||||||| ||
Db 1236 TCCCGTCTGTGCATCCCAGGTGCTGGTATTAAAGGCGAGCTCACCA 1190

RESULT 8
US-08-442-282-4/c
Sequence 4, Application US/08442282
Patent No. 5760204
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

RESULT 7
US-07-757-390-16/c
Sequence 16, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor

```
Qy  4159  tgccctggcactcattctgtgtagacaggctagcctcaaactcacaaacctacctgcctct 4218
          |||||||||||||||||||||||||||||  || |||  |||||||||| || |||||||
Db  3215  TGTCCTGGGACTCACTTTGTAGACCAGGCTGGCCTCCGAACTCAGAAATCTACCTGCCTCT 3156

Qy  4219  gccttcccagtgctggccactaaagatgtggccaacaacaactagta 4264
          |||| |||| ||||||   |||||| | ||| ||| | ||| ||
Db  3155  GCCTCCGAGTGCTGGGATTAAAGGCATGCACCACCACACCCAGCA 3110
```

RESULT 4
US-08-555-723B-3
Sequence 3, Application US/08557234B
Patent No. 5837534

GENERAL INFORMATION:
APPLICANT: OLSON, Eric N.
APPLICANT: LI, Li
APPLICANT: MIANO, Joseph M.
TITLE OF INVENTION: SMOOTH MUSCLE 22' PROMOTER, GENE TRANSFER
TITLE OF INVENTION: VECTORS CONTAINING THE SAME, AND METHOD OF
TITLE OF INVENTION: USE OF THE SAME TO TARGET GENE EXPRESSION IN
TITLE OF INVENTION: ARTERIAL SMOOTH MUSCLE CELLS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 2037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,723B
FILING DATE: 14-NOVEMBER-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6663
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3892 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
HYPOTHETICAL: NO
; US-08-555-723B-3

Query Match   0.8%; Score 100.2; DB 4; Length 3892;
Best Local Similarity 61.9%; Pred. No. 6.1e-15;
Matches 159; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
Qy  4000  tttggctctggtaactgtccaggggcctctctgaccacaggggttgtagcccagcgctcc 4059
          ||||  |||||| |||||||||||||||||||||||||||||      |||     |
Db  3326  TTTCCTTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 3385

Qy  4060  ctctcttctcccccccccctcttcttttgctctggagacttaattttttctttttcctt 4119
          |  ||||  ||||||||||||||||||||||||||    |||||     ||
Db  3386  TCTTTCTTTTCTTTTCTTTCTCTTCTCACTCTCTCTAATTTCTTTCTTTTTTTTTTTTCTT 3445

Qy  4120  ttggctttttgagacaggggtttctctgtacagccctggtcgtcgcctggcactcattctgta 4179
          | ||||  |||||||||| |||| |||||| | |||||   |||||||||
Db  3446  TCTTTTTTTTAGACAGGGTTTCCTCTGTGTGTAGCCCTGGCTGTCGAACTCACTCTTTA 3505
```

```
Qy  4180  gaccaggctagcctcaaactcacaaacctacctgcctcttccaagtgtgctggcacta 4239
          ||||||||| | |||||||| |||  ||| || ||||||| || ||  |||||||||
Db  3506  GACCAGGCTGGCCTGCCTCGAACTCAGAAATCTACCTGCCTCCCAAGTGCTGGGATTA 3565

Qy  4240  aagatgtgggccaccac 4256
          |||  ||||||||||
Db  3566  AAGGCGTGTGCACCAC 3582
```

RESULT 5
US-08-686-878A-38
Sequence 38, Application US/08686878A
Patent No. 5708157

GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: LaValle, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,878A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
; US-08-686-878A-38

Query Match   0.8%; Score 99.8; DE 2; Length 535;
Best Local Similarity 82.5%; Pred. No. 2.6e-15;
Matches 113; Conservative 1; Mismatches 23; Indels 0; Gaps 0;

```
Qy  4103  tttttttctttttttgctttttgaacagggttttctgtacagccctgctgcc 4162
          ||||  ||||||||||||  ||||||||||||||||||| ||  ||||||
Db  399   TTTGTTGGGTTTCTTTTGGTTTTGGAGACAGGGTTTCTCTGTATTAGCCCTGGCTGTC 458

Qy  4163  ctggcactcattctgtagaccaggctagcctcaaactcacaaacctacctgcctctctgcc 4222
          ||:||||||||||||||||||||||||||||||  |||||||||||||||
Db  459   CTARAACTCACTCTGTAGACCAGGCTGGCCTCGACTCAGAAACTCAGAAATCTGCCCTGGCCT 518

Qy  4223  ttccagtgctggcacta 4239
          |||||||||||| ||
Db  519   CCCAAGTGCTGGGATTA 535
```

Db   1606 TGTCCTGGCACTCACTCTGTAGACCACAGGCTCGGCCTCAAACTCAGAAATCTGCCTGCCTCT 1665
            ||||||||| ||||| ||||||| ||||||||||| ||||||| ||| ||||||||||||
Qy   4219 gccttcccagtgctggcactaaagatgtgggccacac 4256
            ||||||| || ||||| | | |||
Db   1666 GCCTCCCAAATGCTGGGATTAAAGGCTTGCACCAGGAC 1703

RESULT 2
US-08-299-849B-5
Sequence 5, Application US/0829984B
Patent No. 5612201

GENERAL INFORMATION:
~ APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
~ APPLICANT: Leth, Bernard; Szikora, Jean-Pierre; De Smet, Charles;
~ APPLICANT: Chomez, Patrick
~ TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
~ TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
~ NUMBER OF SEQUENCES: 48
~ CORRESPONDENCE ADDRESS:
~   ADDRESSEE: Felfe & Lynch
~   STREET: 805 Third Avenue
~   CITY: New York City
~   STATE: New York
~   ZIP: 10022
~ COMPUTER READABLE FORM:
~   MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
~   COMPUTER: IBM
~   OPERATING SYSTEM: PC-DOS
~   SOFTWARE: Wordperfect
~ CURRENT APPLICATION DATA:
~   APPLICATION NUMBER: US/08/299,849B
~   FILING DATE: 1-SEPTEMBER-1994
~   CLASSIFICATION: 435
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: 08/037,230
~   FILING DATE: 26-MARCH-1993
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: PCT/US92/04354
~   FILING DATE: 22-MAY-1992
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: 07/807,043
~   FILING DATE: 12-DECEMBER-1991
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: 07/764,364
~   FILING DATE: 23-SEPTEMBER-1991
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: 07/728,838
~   FILING DATE: 9-JULY-1991
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: 07/705,702
~   FILING DATE: 23-May-1991
~ ATTORNEY/AGENT INFORMATION:
~   NAME: Hanson, No. 561220Iman D.
~   REGISTRATION NUMBER: 30,946
~   REFERENCE/DOCKET NUMBER: LUD 5355
~ TELECOMMUNICATION INFORMATION:
~   TELEPHONE: (212) 688-9200
~   TELEFAX: (212) 838-3884
~ INFORMATION FOR SEQ ID NO: 5:
~ SEQUENCE CHARACTERISTICS:
~   LENGTH: 4698 base pairs
~   TYPE: nucleic acid
~   STRANDEDNESS: single
~   TOPOLOGY: linear
~ MOLECULE TYPE: genomic DNA

Query Match   0.9%; Score 111.6; DB 1; Length 4698;
Best Local Similarity 81.6%; Pred. No. 1.1e-17;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy   4099 ttaattttttctttcttttggcttttgagacaggtgtttctctgtacagccctggc 4158
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1546 TTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTGGAGACAGGGTTTCTCTTTGTATCCCTGGC 1605
            ||||||| |||||| ||| | ||||||| ||||||||| ||||||||| ||||||||||||
Qy   4159 tgccctggcactcattcctgtgaccacaggctagcctcaaaactcacacaaacctaacctgcctct 4218
            ||||||| |||| ||||||||| | ||||||||||||||||| |||| || ||||||||| ||
Db   1606 TGTCCTGGCACTCACTCTGTAGACCACAGGCTCGGCCTCAAACTCAGAAATCTGCCTGCCTCT 1665

Qy   4219 gccttcccagtgctggcactaaagatgtgggccacac 4256
            |||| || ||| |||| ||| | ||| |||
Db   1666 GCCTCCCAAATGCTGGGATTAAAGGCTTGCACCAGGAC 1703

RESULT 3
US-07-923-976-1/c
Sequence 1, Application US/07923976
Patent No. 5574136

GENERAL INFORMATION:
~ APPLICANT: Nagata, Shigekazu
~ APPLICANT: Fukunaga, Rikiro
~ TITLE OF INVENTION: DNA Encoding Granulocyte
~ TITLE OF INVENTION: Colony-Stimulating Factor Receptor
~ NUMBER OF SEQUENCES: 8
~ CORRESPONDENCE ADDRESS:
~   ADDRESSEE: Jones, Tullar & Cooper, P.C.
~   STREET: P.O. Box 2266 Eads Station
~   CITY: Arlington
~   STATE: Virginia
~   ZIP: 22202
~ COMPUTER READABLE FORM:
~   MEDIUM TYPE: Floppy disk
~   COMPUTER: IBM PC compatible
~   OPERATING SYSTEM: PC-DOS/MS-DOS
~   SOFTWARE: PatentIn Release #1.0, Version #1.25
~ CURRENT APPLICATION DATA:
~   APPLICATION NUMBER: US/07/923,976
~   FILING DATE: 19920922
~   CLASSIFICATION: 435
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: JP 74539/1990
~   FILING DATE: 23-MAR-1990
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: JP 176629/1990
~   FILING DATE: 03-JUL-1990
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: PCT/JP91/00375
~   FILING DATE: 22-MAR-1991
~ ATTORNEY/AGENT INFORMATION:
~   NAME: Hellwege, James W.
~   REGISTRATION NUMBER: 28,808
~   REFERENCE/DOCKET NUMBER: 514853
~ TELECOMMUNICATION INFORMATION:
~   TELEPHONE: 703-415-1500
~   TELEFAX: 703-415-1508
~ INFORMATION FOR SEQ ID NO: 1:
~ SEQUENCE CHARACTERISTICS:
~   LENGTH: 3293 base pairs
~   TYPE: NUCLEIC ACID
~   STRANDEDNESS: single
~   TOPOLOGY: linear
~ MOLECULE TYPE: cDNA to mRNA
~ FEATURE:
~   NAME/KEY: CDS
~   LOCATION: 180..2690
US-07-923-976-1

Query Match   0.9%; Score 103.6; DB 1; Length 3293;
Best Local Similarity 76.5%; Pred. No. 8.2e-16;
Matches 127; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy   4099 ttaattttttctttcttttggcttttgagacaggtgtttctctgtacagccctggc 4158
            ||| ||| |||||| ||| || ||||||||| ||||| |||||||||||||||||
Db   3275 TTTATTCAGCTTTGCTTAATTTTTTTTTTTTTTTCAAGACACAGGGTTTCTCTGTGATAGGTCTGGC 3216

Title:
Perfect score: 11832
Sequence: 1 gcggccgtgcagtgattac.............tatgtgctgggtgggggga 11832

Scoring table: IDENTITY_NUC

Searched: 176461 seqs, 4538279 residues

Database :   Issued_Patents_NA:*
                      1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                      2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                      3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
                      4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
                      5: /cgn2_6/ptodata/2/ina/PCTUS9_COMB.seq:*
                      6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Run on: September 19, 1999, 11:39:06 ; Search time 125.52 Seconds
                      (without alignments)
                      8641.786 Million cell updates/sec

OM nucleic - nucleic search, using sw model

                      Copyright (c) 1993 - 1998 Compugen Ltd.

                      GenCore version 4.5

US-09-037-657-38

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	111.6	0.9	4698	1	US-07-807-043B-5	Sequence 5, Appl1
2	111.6	0.9	4698	1	US-08-299-849B-5	Sequence 5, Appl1
3	103.6	0.9	3293	1	US-07-923-976-1	Sequence 1, Appl1
4	100.2	0.8	3892	4	US-08-555-723B-3	Sequence 3, Appl1
5	99.8	0.8	535	2	US-08-686-878A-38	Sequence 38, Appl1
6	99	0.8	1355	1	US-07-757-390-4	Sequence 4, Appl1
7	99	0.8	1355	1	US-07-757-390-16	Sequence 16, Appl1
8	99	0.8	1355	3	US-08-442-282-4	Sequence 4, Appl1
9	99	0.8	1355	3	US-08-442-282-16	Sequence 16, Appl1
10	99	0.8	1355	3	US-08-442-281-4	Sequence 4, Appl1
11	99	0.8	1355	3	US-08-442-281-16	Sequence 16, Appl1
12	95	0.8	7218	2	US-08-232-463-14	Sequence 14, Appl1
13	94.2	0.8	1808	1	US-07-757-390-2	Sequence 2, Appl1
14	94.2	0.8	1808	1	US-07-757-390-15	Sequence 15, Appl1
15	94.2	0.8	1808	3	US-08-442-282-2	Sequence 2, Appl1
16	94.2	0.8	1808	3	US-08-442-282-15	Sequence 15, Appl1
17	94.2	0.8	1808	3	US-08-442-281-2	Sequence 2, Appl1
18	94.2	0.8	1808	3	US-08-442-281-15	Sequence 15, Appl1
19	90.6	0.8	240	1	US-08-487-748A-6	Sequence 6, Appl1
20	89.6	0.8	2797	2	US-08-555-723B-4	Sequence 4, Appl1
21	87.6	0.7	5687	4	US-08-380-403A-3	Sequence 3, Appl1
22	86.6	0.7	7218	2	US-08-232-463-14	Sequence 14, Appl1
23	85	0.7	3383	5	PCT-US95-09098-1	Sequence 1, Appl1
24	84.8	0.7	1753	5	PCT-US95-02576-6	Sequence 6, Appl1
25	82	0.7	26700	2	US-08-472-217-1	Sequence 1, Appl1
26	82	0.7	26700	1	US-08-488-199-5	Sequence 5, Appl1
27	81.2	0.7	4016	5	PCT-US95-08354A-1	Sequence 1, Appl1
28	79.6	0.7	1179	3	PCT-US95-794-4	Sequence 4, Appl1
29	77.4	0.7	3963	1	US-08-464-961-1	Sequence 1, Appl1
30	77.4	0.7	3963	5	PCT-US96-0833-1	Sequence 1, Appl1
31	77	4	5253	4	US-08-290-731C-3	Sequence 3, Appl1
32	74.8	0.5	3288	1	US-08-620-694A-1	Sequence 1, Appl1
33	73.6	0.5	200	1	US-08-222-177A-8	Sequence 8, Appl1
34	73.4	0.5	1919	1	US-07-991-587A-1	Sequence 1, Appl1
35	73.4	1	1919	1	US-08-309-985-1	Sequence 1, Appl1
36	73.4	0.5	2738	4	US-08-795-868-17	Sequence 17, Appl1
37	72.6	0.6	4157	4	US-08-871-266B-1	Sequence 1, Appl1
38	72.2	0.6	5191	2	US-08-340-428B-1	Sequence 1, Appl1
39	72.2	0.6	5191	5	PCT-US93-07306-1	Sequence 1, Appl1
40	71.6	0.6	2997	5	PCT-US95-07349-3	Sequence 3, Appl1
41	70.8	0.6	2757	4	US-08-627-254C-26	Sequence 26, Appl1
42	70.4	0.6	91	1	US-08-222-177A-107	Sequence 107, Appl1
43	70	0.6	10614	1	US-08-135-511-35	Sequence 35, Appl1
44	70	0.6	10614	3	US-08-187-453-35	Sequence 35, Appl1
45	67.6	0.6	1888	5	PCT-US95-02576-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-07-807-043B-5
Sequence 5, Application US/07807043B
Patent No. 5342774
GENERAL INFORMATION:
APPLICANT: Boon, Thierry, Van den Eynde, Beno t
TITLE OF INVENTION: Tumor Rejection Antigen Precursors, Tumor
TITLE OF INVENTION: Rejection Antigens and Uses Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 ZIP: 10022
COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/807,043B
 FILING DATE: 19911212
 CLASSIFICATION: 424
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/764,364
 FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/728,838
 FILING DATE: 9-JULY-1991
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/705,702
 FILING DATE: 23-May-1991
ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 534277.4man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 253.3
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
 LENGTH: 4698 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: singular
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
US-07-807-043B-5

Query Match   0.98; Score 111.6; DB 1; Length 4698;
Best Local Similarity   81.6%; Pred. No. 1.1e-17;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy   4099 ttaatttttctttctttttgcttttttgagacaggttttctcgtacagccctggc 4158
          || |||||| ||| ||||| |||| |||||||||||||||||||||||||||||||
Db   1546 TTTTTTTTTTTTTTTTTTTTTTTGGTTTTTTCGAGACAGGGTTTCTCTGTTGTATCCCTGGC 1605

Qy   4159 tgccctggcactcattcctgtagacccaggctagcctcaaactcacaaactaccctgcctct 4218
          ||||||||||||||||||||| ||||||||||||||||||||| ||| |||||||||||-
Db   4159 tgccctggcactcattcctgtagacccaggctagcctcaaactcacaaactaccctgcctct 4218

TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pHzgpt-F1s
US-08-232-463-14

Query Match
Best Local Similarity 3.6%; Score 95; DB 2; Length 7218;
Matches 14; Conservative 254; Pred. No. 1.6e-13; Mismatches 119; Indels 0; Gaps 0;
US-08-232-463-14

```
Qy  3769  ctggggtgtgggtgcagcctctgttctccgagcctcttggaaatcggtttttgtttttgtt  3828
                ||||  ||  ||| | |:::::::::|::::::::::YYYYYYYYYYYYYYYYYYY
Db  1043  CTGCAGGTCGAGGGAGCTTGGCATYYYYYYYYY::::::YYYYYYYYYYYYYYYYYYY  1102

Qy  3829  tttgtttttttccaatactctttttccctcatccaatccaccccggactgttttcctccctaag  3888
          YYYYYYYYYYYYYY:::::::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db  1103  YYYYYYYYYYYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYYYYYYYYYYYYY  1162

Qy  3889  ggttgagagcccctgcgtctccctaaccttttcttgcttctaccccaggggccttgca  3948
          YYYYYYYYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYYYYYYYYYYY
Db  1163  YYYYYYYYYYYYYYYYYYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYYY  1222

Qy  3949  catgagtccaccctcccccccttgcccaactgggcttcactgtcgcattggctct  4008
          YYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db  1223  YYYYYYYYYYYYYYYYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYY  1282

Qy  4009  tggtaactgtcccaggccctcctgacacacaggggttgtagcccccagctccctcttct  4068
          YYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db  1283  YYYYYYYYYYYYYYYYYYYYYYYYYYY::::::YYYYYYYYYYYYYYYY  1342

Qy  4069  cctcccccttctcttttgcttctgagacttaattaatttttctttcttttggcttt  4128
          YYYYYYYYYYYYY::::::::::YYYYYYYYYYYYYYYYYYYYYYYYYYYY
Db  1343  YYYYYYYYYYYYYYYYYYYYYYYY::::::::::YYYYYYYYYYYYYYYY  1402

Qy  4129  tgagacaagggtttctctgtacaagccct  4155
          YYYYYYYYYYYYYY  ||||
Db  1403  YYYYYYYYYYYYYYYYYYYYYY  1429
```

RESULT 13
US-07-757-390-2/c
Sequence 2, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1808 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-757-390-2

Query Match
Best Local Similarity 0.8%; Score 94.2; DB 1; Length 1808;
Matches 122; Conservative 0; Pred. No. 1.2e-13; Mismatches 28; Indels 1; Gaps 1;
US-07-757-390-2

```
Qy  5549  gtttggttgagttttgttattgagacaggcggctctctgtgtagtcctggctcctctt  5608
          |||||  ||  | | | | || || || ||||||||||||||  ||||||||||  |
Db  1808  GTTTGTTGTTTGTTTGTTGTTTTTTGGAGACAGGGTTTCTCTGTGTAGCCTGGCTGTCC-T  1750

Qy  5609  ggaactcactctgtagaccaggctggcctttgaactcagaaaatccgcctcgcctgctcc  5668
          ||||||||| | |||||||| ||||||| |||||||||||||| |||| ||||  |  ||
Db  1749  GGAACTCACTTTATAGACCAAGCTGGCCTTGAACTCAGAAAATCCTCCCGCCTCGTGTGCATCC  1690

Qy  5669  caagtgctgttagataaaggtgtgcactgca  5699
          ||  |||||  ||||  ||||||| || |||
Db  1689  CAGGTGCTGGGTATTAAAAGCGCAGCTCTACCA  1659
```

RESULT 14
US-07-757-390-15/c
Sequence 15, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tominaga, Akira
APPLICANT: Takagi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1808 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 303..1547

US-07-757-390-15

```
Query Match        0.8%;  Score 94.2; DB 1; Length 1808;
Best Local Similarity  80.8%;  Pred. No. 1.2e-13;
Matches 122; Conservative   0; Mismatches  28; Indels   1; Gaps   1;

Qy   5549 gtttggtttgagtttgttcattttgagacagggggctcttctgttagtcctggctgttcctt 5608
          ||||||| |||||||||| || ||||| ||||||||||||||||||| |||||||||||||
Db   1808 GTTTGTTTGTTTGTTGTTGTTTTTTCGAGACAGGGTTTCTGTGTAGCCCTGGCTGTCC-T 1750

Qy   5609 ggaactcactcctgtagaccaggctgacctcagaaatccgcctgtcttgtgctcc 5668
          ||||||||||||||||||| |||||||||||||||||||| |||| ||
Db   1749 GGAACTCACTTTATAGACCAAGCTGGCCTCGGACTCAGAAATCCTCCCGTCTGTGGCATCC 1690

Qy   5669 caagtgcttagattaaaggtgtgcactgcca 5699
          ||||||| |||||||||| | |||| ||| |||
Db   1689 CAGGTGCTGGGTATTAAAGGCGAGCTCTACCA 1659
```

RESULT 15
US-08-442-282-2/c
; Sequence 2, Application US/08442282
; Patent No. 5760204
GENERAL INFORMATION:
  APPLICANT: Takatsu, Kiyoshi
  APPLICANT: Tominaga, Akira
  APPLICANT: Takagi, Satoshi
  APPLICANT: Murata, Yoshiyuki
  TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
  NUMBER OF SEQUENCES: 18
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Pennie & Edmonds
    STREET: 1155 Avenue of the Americas
    CITY: New York
    STATE: New York
    COUNTRY: U.S.A.
    ZIP: 10036-2711
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/442,282
    FILING DATE:
    CLASSIFICATION: 536
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/757,390
    FILING DATE: 10-SEP-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: Misrock, S. Leslie
    REGISTRATION NUMBER: 18,872
    REFERENCE/DOCKET NUMBER: 7005-030
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 212 790-9090
    TELEFAX: 212 869864/9741
    TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1808 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
US-08-442-282-2

```
Query Match        0.8%;  Score 94.2; DB 3; Length 1808;
Best Local Similarity  80.8%;  Pred. No. 1.2e-13;
Matches 122; Conservative   0; Mismatches  28; Indels   1; Gaps   1;

Qy   5549 gtttggtttgagtttgttcattttgagacagggggctcttctgtgtagtcctggctgtccctt 5608
```

```
Db   1808 GTTTGTTTGTTTGTTGTTGTTTTTTCGAGACAGGGTTTCTGTGTAGCCCTGGCTGTCC-T 1750

Qy   5609 ggaactcactcctgtagaccaggctgacctcagaaatccgcctgtcttgtgctcc 5668
          ||||||||||||||||||| |||||||||||||||||||| |||| ||
Db   1749 GGAACTCACTTTATAGACCAAGCTGGCCTCGGACTCAGAAATCCTCCCGTCTGTGGCATCC 1690

Qy   5669 caagtgcttagattaaaggtgtgcactgcca 5699
          ||||||| |||||||||| | |||| ||| |||
Db   1689 CAGGTGCTGGGTATTAAAGGCGAGCTCTACCA 1659
```

Search completed: September 19, 1999, 16:09:05
Job time: 16199 sec

Left column (RESULT 11)

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797892.
  Contact: Robert Strausberg, Ph.D.
  Tel: (301) 496-1550
  Email: Robert_Strausberg@nih.gov
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1615 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence: stop: 407.

FEATURES Location/Qualifiers
 source 1..431
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="4p16.1-4pter"
  /clone="IMAGE:1930606"
  /clone_lib="Soares_fetal_lung_NbHL19W"
  /dev_stage="19 weeks"
  /lab_host="DH10B (ampicillin resistant)"
  /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
  modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
  strand cDNA was primed with a Not I - oligo(dT) primer
  [5'-TGTTACCAATTCGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTTTT-3'],
  double-stranded cDNA was size selected, ligated to Eco RI
  adapters (Pharmacia), digested with Not I and cloned into
  the Not I and Eco RI sites of a modified pT7T3 vector
  (Pharmacia). Library went through one round of
  normalization to a Cot = 5. Library constructed by Bento
  Soares and M.Fatima Bonaldo. This library was constructed
  from the same fetus as the fetal heart library, Soares
  fetal heart NbHH19W."

BASE COUNT 68 a 169 c 122 g 72 t
ORIGIN

Query Match 1.9%; Score 219.2; DB 45; Length 431;
Best Local Similarity 86.4%; Pred. No. 1.7e-37;
Matches 242; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
Qy 6331 cacagctgtaatcagccccaggagcccacccacccttctcatggctctccctgcaagctac 6390
        ||||||  ||||||||||||||||| ||||||| ||||||||||||||||| || ||
Db   58 CACACGTGATCAGTCCCCAGGAGATCCCACCACGCTTCATCGGGCTCTCCCTGCGGCCAC 117

Qy 6391 ctgctcttatacatggagacacacctggggccacccgctgagggggctccactgaccttca 6450
        ||||| || |||||||||  | |||||||||||||||||||||| |||||||| ||||||
Db  118 CTGCTCAGTGCACGGAGGACCCACCACCAGGAGCCGGCCCTCACTGGACCCTCAA 177

Qy 6451 tggtcgccgccctgacctgtcccggcctccttaacactcacacctggccctggc 6510
        ||||| ||  || ||||||| ||||||||||||||||||||||||||| |||
Db  178 CGGGCGCGCCCCCTGAGCTCTCCCGGTACTCAACGCTCACCTGGCTCTGGC 237

Qy 6511 cctggctaaccttaatgtggtccaggcagcagcagagagacaaatctggttgttcaagccg 6570
        |||||| ||||| |||| |||||||||||||| |||||  |||||||||||||||||||
Db  238 CCTGGCCCAACCTCAAATGGGTCCAGGCAGCAGCACCTCGTGTGCCACCGCCCG 297

Qy 6571 agacggcagcagcattctggctggctcctggcctcctctatatgttggc 6610
        |||||||||||| |||||||| |||||||||||||||||||||||||
Db  298 TGACGGCAGCAGCAATCCTGGCTCCTGCCTCCTTATAGTTGGC 337
```

RESULT 11
LOCUS AI670108 385 bp mRNA EST 14-MAY-1999
DEFINITION we65103.x1 Soares_thymus_NHFTh Homo sapiens cDNA clone
  IMAGE:2345981 3' similar to TR:O75462 O75462 CYTOKINE-LIKE FACTOR-1
  PRECURSOR ;, mRNA sequence.
ACCESSION AI670108
NID g4834882
VERSION AI670108.1 GI:4834882
KEYWORDS EST.

Right column (RESULT 12)

SOURCE human.
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 385)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 9, 1996 this sequence version replaced gi:1133091.
  Contact: Robert Strausberg, Ph.D.
  Tel: (301) 496-1550
  Email: Robert_Strausberg@nih.gov
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40UP from Gibco
  High quality sequence: stop: 372.

FEATURES Location/Qualifiers
 source 1..385
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="2"
  /clone="IMAGE:2345981"
  /clone_lib="Soares_thymus_NHFTh"
  /dev_stage="fetal"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: thymus, pooled; Vector: pT7T3D-Pac
  (Pharmacia) with a modified polylinker; Site_1: Not I;
  Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
  oligo(dT) primer [5'-
  TGTTACCAATCGAAGTGAGAGCGGCCGCAAGGTTTTTTTTTTTTTTTTT 3'],
  double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not I
  and Eco RI sites of the modified pT7T3 vector. Library
  went through one round of normalization. Library
  constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 55 a 156 c 108 g 66 t
ORIGIN

Query Match 1.8%; Score 215.4; DB 50; Length 385;
Best Local Similarity 85.4%; Pred. No. 1.1e-36;
Matches 240; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy 6330 acacagctgtaatcagccccaggacccccacccacccttctcatcggctccctgcaagcta 6389
        |||||||||||||||||||||||  |||||||||||| ||||  ||||||||||||||||
Db   75 ACACAGCTGTGATCAGTCCCCAGGATCCCACCACGCTTCTCATCGGGCTCTCCCTGCGGCCA 134

Qy 6390 cctgctctatacatggagacacacctggggccaccgctgagggggctctactggaccttca 6449
        ||||||| ||||||||||| |||||||||||||| ||||||||||||||||||| ||||
Db  135 CCTGCTCAGTGCACGGAGGACCCACCACCAGGAGCCGGCCCTCACTGGACCCTCA 194

Qy 6450 atggtcgccgccctgacctgtcccggcctccttatcacctcaccctggcctggg 6509
        | |  ||| || ||||||| |||||||||||| |  |||||||||||||| ||
Db  195 ACGGGCGCGCCCCCTGAGCTCTCCCGTGTACTCAACGCCTCACCCTGGCTCTGG 254

Qy 6510 ccctggctaaccttaatggtccaggcagcagcagagagacaaatctggttgttcaagccc 6569
        |||||||| |||| | |||||||||||||||||||  |||||||||||||||||| ||
Db  255 CCCTGGCCCAACCTCAAATGGGTCCAGGCAGCAGCACCTCGTGTGGGGGACAACCTCGTTGTGCCCACCCCC 314

Qy 6570 gagacggcagcagcattctggctggctcctggcctcctctatatgttggc 6610
        |||||||||||| ||||||||| |||||||||||||||||||||||||
Db  315 GTGACGGCAGCAGCAATCCTGGCTCCTGGCTCCTTATAGTTGGC 355
```

RESULT 12
LOCUS W66776/c 482 bp mRNA EST 14-JUN-1996
DEFINITION me17b11.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
  clone IMAGE:387741 5' similar to PIR:B38252 B38252 granulocyte
  colony-stimulating factor receptor precursor ;, mRNA sequence.
ACCESSION W66776

```
NID         g1375694
VERSION     W66776.1    GI:1375694
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 482)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterson,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:785250.

            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            MGI:239573
            Seq primer: ETPrimer
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1..482
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="10 pter-cen"
                     /clone="IMAGE:387741"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc embryo"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGCAGGAGGCGGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
                     State Univ. from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.
                     "
BASE COUNT       96 a    138 c    147 g    101 t
ORIGIN
```

```
Qy     8604  gtgacaacggacccccccaccccgacgtgcacgtggggcctggaggaccag 8663
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      317  GTGACCACGGAGACCCCCACACCCGACGTGCACGTGCAGCCCGTTGCGGGGGCCTGGAGGACCAG 258
```

```
Qy     8664  ctgagtgtgcgctggctcaccacacacacagcctccaagacctccaagtac 8723
             ||| ||||||||||||||||||||||||||||||| ||||||||||||||
Db      257  CTGAGTGTGCGCTGGGTCTCACCACACACCAGCTCTCAAGGAATTTCCTCTTCAAGCCAAGTAC 198
```

```
Qy     8724  cag-atccgtcaccgctgggaggacagcggtgaagtgcccgtcccgcccggac 8782
             ||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db      197  CAGTAATTCGCTACCGCCTGGGAGGACACGGTGGACTGGA------------------ 159
```

```
Query Match       1.8%;  Best Local Similarity  81.1%;
Matches 305; Conservative   0; Mismatches   2; Indels  69;  Gaps   2;
                  Score 214.8;  DB 26;   Length 482;   Pred. No. 1.6e-36;
```

```
RESULT 13
AV032198
LOCUS       AV032198             262 bp    mRNA              EST       03-JUN-1999
DEFINITION  AV032198 Mus musculus adult C57BL/6J cerebellum Mus musculus cDNA
            clone 1500034F20, mRNA sequence.
ACCESSION   AV032198
NID         g4831746
VERSION     AV032198.1  GI:4831746
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 262)
AUTHORS     Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Ito,M., Aizawa,K.,
            Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
            Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
            Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
            Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
            Sugahara,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y.,
            Tominaga,N., Watanabe,S., Yagane,M., Yamamura,T., Yokota,T.,
            Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
            RIKEN Mouse ESTs
TITLE       RIKEN Mouse ESTs
JOURNAL     Unpublished (1999)
COMMENT     On Jun 22, 1998 this sequence version replaced gi:3246782.

            Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-res@rtc.riken.go.jp
            This clone can be obtained from the RIKEN Mouse cDNA
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length cDNA
            (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1..262
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="1500034F20"
                     /clone_lib="Mus musculus adult C57BL/6J cerebellum"
                     /sex="male"
                     /tissue_type="cerebellum"
                     /dev_stage="adult"
BASE COUNT       63 a    63 c    69 g    67 t
ORIGIN
```

```
Qy     8783  ccgccccctgacccgccgcccccgccatctgactcctcccctcaccgtgcaggtgggtggatgac 8842
             |||||||||||||||||||||||||||||||||||||||||||||
Db      159  ------------------------------------------------AGGTGGTGGATGAC 146
```

```
Qy     8843  gtcagcaaccacagacctcctgccgtctccgcggggcctgaagaccggcgaccgtttacttcgtc 8902
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      145  GTCACCAACCACAGACCTCCTCGCCGTCTCCGCGGGGCCTGAAGCCCGGCGACCGTTTACTTCGTC 86
```

```
Qy     8903  caagtgcgttgtaacccccattcggggatctatgggtcgaaaaagggggagatcttgagcggag 8962
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85   CAAGTGCGTTGTAACCCATTCGGGGGATCTATGGGGTCGAAAAAGGGCGGGAATCTGGAGCGAG 26
```

```
Qy     8963  tggagcaccacccacg 8978
             ||||||||||||| ||
Db      25   TGGAGCCACCCCATCG 10
```

```
Query Match       1.8%;  Score 213;  DB 50;   Length 262;
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 1998 Compugen Ltd.

GenCore version 4.5

Run on: September 19, 1999, 11:37:06 ; Search time 1717.66 Seconds
(without alignments)
13587.681 Million cell updates/sec

Title: US-09-037-657-38
Perfect score: 11832
Sequence: 1 gcggccgctgcagtgattac.......tatgtgctggggtggggggga 11832
Database :   2546578 seqs, 986266752 residues
Searched: IDENTITY_NUC
Scoring table:

EST:*

No.	Label		No.	Label
1:	em_est1:*		30:	gb_est9:*
2:	em_est2:*		31:	gb_est10:*
3:	em_est3:*		32:	gb_est11:*
4:	em_est4:*		33:	gb_est12:*
5:	em_est5:*		34:	gb_est13:*
6:	em_est6:*		35:	gb_est14:*
7:	em_est7:*		36:	gb_est15:*
8:	em_est8:*		37:	gb_est16:*
9:	em_est9:*		38:	gb_est17:*
10:	em_est10:*		39:	gb_est18:*
11:	em_est11:*		40:	gb_est19:*
12:	em_est12:*		41:	gb_est20:*
13:	em_est13:*		42:	gb_est21:*
14:	em_est14:*		43:	gb_est22:*
15:	em_est15:*		44:	gb_est23:*
16:	em_est16:*		45:	gb_est24:*
17:	em_est17:*		46:	gb_est25:*
18:	em_est18:*		47:	gb_est26:*
19:	em_est19:*		48:	gb_est27:*
20:	em_est20:*		49:	gb_est28:*
21:	em_est21:*		50:	gb_est29:*
22:	gb_est1:*		51:	gb_est30:*
23:	gb_est2:*		52:	gb_est31:*
24:	gb_est3:*		53:	em_est21:*
25:	gb_est4:*		54:	em_est22:*
26:	gb_est5:*		55:	em_est23:*
27:	gb_est6:*		56:	em_est24:*
28:	gb_est7:*		57:	em_est25:*
29:	gb_est8:*		58:	em_est26:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

c	Result No.	Score	% Query Match	Length	DB ID		Description
c	1	248	2.1	445	27	AA039053	mi99d07.r
c	2	221.8	1.9	462	45	AI394468	tf79dl2.x
	3	221.8	1.9	474	46	AI421423	tf25h01.x
	4	220.2	1.9	466	43	AI161002	qb69g04.x
	5	220.2	1.9	466	43	AI185780	qe44h04.x
	6	220.2	1.9	477	43	AI185924	qe50c05.x
	7	220.2	1.9	447	43	AI187074	qe38a03.s
	8	220.2	1.9	417	48	AI269388	qi26b05.x
	9	219.6	1.9	332	48	AI579568	UI-R-G0-u
	10	215.4	1.9	431	45	AI333812	qp93e12.x
	11	215.4	1.8	385	50	AI670108	we65f03.x
	12	214.8	1.8	482	26	W66776	me17b11.r
	13	213	1.8	262	50	AV032198	AV032198
	14	207.2	1.8	319	48	AI574687	UI-R-G0-u
c	15	200.2	1.7	410	27	AA042914	zk56f01.s
	16	195.6	1.7	390	26	W17583	mb75b01.r1
	17	176.8	1.5	273	23	R87407	yn88d09.s1
c	18	174.8	1.5	464	29	AA049280	mj45d02.r
	19	171	1.4	428	45	AA866388	UI-R-A0-a
	20	169.4	1.4	229	49	AV010798	AV010798
c	21	164	1.4	503	29	AA049278	mj45c04.r
c	22	160.4	1.4	465	26	W46604	zc32h10.s1
	23	159.2	1.3	455	46	AA925924	UI-R-A1-e
c	24	151	1.3	470	43	AI233311	EST22999
	25	143.8	1.2	167	49	AV011309	AV011309
c	26	135.8	1.1	464	41	AI071408	UI-R-C2-n
c	27	134.8	1.1	235	48	AI575060	UI-R-G0-u
	28	126.8	1.1	496	44	AI316848	u135a03.y
c	29	126.4	1.1	370	48	AI605946	vw35e05.x
c	30	124.8	1.1	397	33	AA390114	vb28f01.r
c	31	124	1.0	422	46	AI451475	mu42a09.x
	32	122.4	1.0	365	49	AI661062	mz81e09.x
c	33	122	1.0	476	36	C79500	C79500 Mous
c	34	122	1.0	482	47	AI509657	vx18g08.y
c	35	121.6	1.0	401	47	AI465800	vw98e12.y
c	36	121.4	1.0	458	27	AA043001	zk56f01.y
c	37	121.4	1.0	587	36	C79507	C79507 Mous
c	38	121.4	1.0	309	48	AI597540	vd75b01.y
c	39	121	1.0	422	37	AA683670	vr03e10.r
c	40	120.8	1.0	800	50	AU067650	AU067650
c	41	120.4	1.0	401	39	AA881428	vx18g08.r
c	42	119.8	1.0	321	46	C79084	C79084 Mous
c	43	119	1.0	398	29	AA153045	mq53g08.r
c	44	119	1.0	389	46	AI428461	mo30f09.x
c	45	118.8	1.0	394	30	AA239691	mv23h01.r

ALIGNMENTS

RESULT 1
AA039053/c
LOCUS   AA039053   445 bp   mRNA   EST   29-AUG-1996
DEFINITION   mi99d07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:474733 5', mRNA sequence.
ACCESSION   AA039053
NID   g1514788
VERSION   AA039053.1   GI:1514788

KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM   Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 445)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
  Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
  Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
  Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
  Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
  WashU-HHMI Mouse EST Project
  Washington University School of MedicineP
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: mouseest@watson.wustl.edu
  This clone is available royalty-free through LLNL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  MGI:285477
  Seq primer: -28M13 rev2 from Amersham
  High quality sequence stop: 441.
FEATURES   Location/Qualifiers
  source   1..445
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="IMAGE:474733"
  /clone_lib="Soares mouse embryo NbME13.5 14.5"
  /sex="unknown"
  /tissue_type="embryo"
  /dev_stage="13.5-14.5dpc total fetus"
  /lab_host="DH10B"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
  polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
  was primed with a Not I - oligo(dT) primer [5'
  TGTTACCAATCTGAAGTGGGAGCGGCCGCGGGAAAATTTTTTTTTTTTTTTTTTTTT
  T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
  14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
  State Univ., from 2]; double-stranded cDNA was ligated to
  Eco RI adaptors (Pharmacia), digested with Not I and
  cloned into the Not I and Eco RI sites of the modified
  pT7T3 vector. Library went through one round of
  normalization, and was constructed by Bento Soares and
  M.Fatima Bonaldo. "
BASE COUNT   85 a   129 c   124 g   107 t
ORIGIN

Query Match   2.1%;   Score 248;   DB 27;   Length 445;
Best Local Similarity   99.6%;   Pred. No. 9.4e-44;
Matches 259;   Conservative   0;   Mismatches   1;   Indels   1;   Gaps   1;

Qy   11534   aggtcctgcgcggctaaacctaaggataggcccatcctcctggtagacctggaggc 11593
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265   AGGTCCTGCGCGGCTAAACCTAAGGATAGGCCCATCCTCCTGGTCAGACCTGGAGGC 206

Qy   11594   tcacctgaattggacccccctcctgtacactcttgggaacaaaagaaaacctaccagggct-g 11652
  |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   205   TCACCTGAATTGGACCCCTCCTGTACCACTCTGGGCAACAAAAGAAAACCTACCAGAGGCTGG 146

Qy   11653   ggcacaatgagctcccacaacacagcttggtccacatgatggtcacacttggatatac 11712
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145   GGCACAATGAGCTCCCACAACCACAGCTTTGGTCCACATGATGGTCACACTTGGATATAC 86

Qy   11713   cccagtgtgtgggtaggtttggggtgtattgcaggggcctcccaagagtcctcttaaataa 11772
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85   CCCAGTGTGTGGGTAGGGTTAGGGTGTATTGCAGGGGCCTCCCAAGAGTCTCTTTAAATAAATAA 26

Qy   11773   aggagttgttcaggtcccga 11792
  ||||||||||||||||||||
Db   25   AGGAGTTGTTCAGGTCCCGA 6

RESULT   2
LOCUS   AI394468   462 bp   mRNA   EST   30-MAR-1999
DEFINITION   tf79d12.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2105495 3',
  mRNA sequence.
ACCESSION   AI394468
NID   g4224015
VERSION   AI394468.1   GI:4224015
KEYWORDS   EST.
SOURCE   human.
  ORGANISM   Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 462)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute / National Institute of Neurological
  Disorders and Stroke, Brain Tumor Genome Anatomy Project
  (CGAP/BTGAP) Tumor Gene Index
JOURNAL   Unpublished (1998)
COMMENT   On Feb 17, 1998 this sequence version replaced gi:2887603.

  Contact: Robert Strausberg, Ph.D.
  Tel: (301) 496-1550
  Email: Robert_Strausberg@nih.gov
  Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
  Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www.bio.llnl.gov/bbrp/image/image.html
FEATURES   Location/Qualifiers
  source   1..462
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="10; 1"
  /clone="IMAGE:2105495"
  /clone_lib="NCI_CGAP_Brn23"
  /tissue_type="glioblastoma (pooled)"
  /lab_host="DH10B"
  /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
  modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
  strand cDNA was primed with a Not I - oligo(dT) primer [5'
  TGTTACCAATCTGAAGTGGGAGCGGCCGCGGGAATTCTTTTTTTTTTTTTTTTTTTT
  T 3']; double-stranded cDNA was ligated to Eco RI
  adaptors (Pharmacia), digested with Not I and cloned into
  the Not I and Eco RI sites of the modified pT7T3 vector.
  Library is normalized, and was constructed by Bento
  Soares and M.Fatima Bonaldo."
  Insert Length: 1631   Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 454.
BASE COUNT   79 a   181 c   122 g   79 t   1 others
ORIGIN

Query Match   1.9%;   Score 221.8;   DB 45;   Length 462;
Best Local Similarity   86.8%;   Pred. No. 4.7e-38;
Matches 244;   Conservative   0;   Mismatches   37;   Indels   0;   Gaps   0;

Qy   6330   acacagctgtaatcagccccccaggagccccaaccctctcatcggctcctccctgcaagcta 6389
  |||||||||| ||| |||||| || |||||||||||||||||||||||||||||| |||
Db   54   ACACAGCTGTGATCAGTCCCCAGGAATCCCACGCTTCTCATCGGCTCCTCCCTGCTGGCCA 113

Qy   6390   cctgctctatacatggagacacctgggcccacgctgagggctctcactggacctca 6449

Db 114 CCTGCTCAGTGACGCTGCACGGAGACCACCACGCGCCGAGGGCGCCTCTACTGGACCCTCA 173
   ||||||||| |||||||||| |||||||| ||||||||||||| ||||||||||||||||||

Qy 6450 atggtcgcgcgccctgagctgtccgcgcctccttaacacctccaccctggccctgg 6509
   ||||||||||||||| ||||| |||||||||||||||||||||||||| |||||

Db 174 ATGGGCGCCGCGCCCCTGCCCCTGACGCTCTCCCGTGTACTCAACGCCTTCCACCTGGCCTCTGG 233

Qy 6510 ccctggctaaccttaatggggtccaggcaggcagtcaggagacaatctggtgtgtcacgccc 6569
   |||||||||| |||| ||||||||||||||||||| |||| |||||||||||||||| ||||

Db 234 CCCTGGCGCCAAACCTCAATGGGTCCAGGCAGGCAGGGGGACAACCTCGTGTGTGCCACGCCC 293

Qy 6570 gagacggccagcattctggcgtgtcctgcctctatgtgtggc 6610
   ||||||||||| |||||||||||||||||||||||||||||||

Db 294 GTGACGGCGAGCCATCCTGGCTGGCTGCCTCCTGCCTCTATGTTGGC 334

RESULT 3

LOCUS AI421423 474 bp mRNA EST 30-MAR-1999

DEFINITION tf25h01.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2097265 3', similar to SW:IL6B_MOUSE O00560 INTERLEUKIN-6 RECEPTOR BETA CHAIN PRECURSOR ;; mRNA sequence.

ACCESSION AI421423

NID   g4267354

VERSION AI421423.1 GI:4267354

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
   Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 474)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1998)

COMMENT On Apr 7, 1998 this sequence version replaced gi:3034955.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

FEATURES Location/Qualifiers
source 1..474
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_lib="NCI_CGAP_Brn23"
   /clone="IMAGE:2097265"
   /tissue_type="glioblastoma (pooled)"
   /lab_host="DH10B"
   /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I - Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCGCGCAATATCTTTTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

Insert Length: 1664 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 450.

  81 a 181 c 128 g 83 t 1 others

Query Match 1.9%; Score 221.8; D3 46; Length 474;
Best Local Similarity 86.8%; Pred. No. 4.7e-38;
Matches 244; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 6330 acacagctgtaatcaagccccaggaccccacccttcctcatcggctcctccctgcaagcta 6389
   ||||| ||||||||| ||||||||||||||||| | |||| |||| |||||||||| ||||

Db 54 ACACAGCTGTGATCAGTCCCCAGGAANCCCACGCGTTCTCATCGGCTCCTCCCTGCGCCCA 113

Qy 6390 cctgtctatacacatggagacaacctggggccacgctgaaggggctctactggaccttca 6449
   |||||||| |||| ||||||||||||||||||||| || ||||||| |||||||||||||||

Db 114 CCTGCTCAGTGACGCTGCACGGAGACCACCACGCGCCGAGGGCGCCTCTACTGGACCCTCA 173

Qy 6450 atggtcgcgcgccctgagctgtccgcgcctccttaacacctccaccctggccctgg 6509
   ||||||||||||||| ||||| |||||||||||||||||||||||||| |||||

Db 174 ATGGGCGCCGCGCCCCTGCCCCTGACGCTCTCCCGTGTACTCAACGCCTTCCACCTGGCCTCTGG 233

Qy 6510 ccctggctaaccttaatggggtccaggcaggcagtcaggagacaatctggtgtgtcacgccc 6569
   |||||||||| |||| |||||||||||||||||| |||| |||||||||||||||| ||||

Db 234 CCCTGGCGCCAAACCTCAATGGGTCCAGGCAGGCAGGGGGACAACCTCGTGTGTGCCACGCCC 293

Qy 6570 gagacggccagcattctggcgtgtcctgcctctatgtgtggc 6610
   ||||||||||| |||||||||||||||||||||||||||||||

Db 294 GTGACGGCGAGCCATCCTGGCTGGCTGCCTCCTGCCTCTATGTTGGC 334

RESULT 4

LOCUS AI161002 466 bp mRNA EST 26-OCT-1998

DEFINITION qb69g04.x1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:1705398 3', mRNA sequence.

ACCESSION AI161002

NID   g3694307

VERSION AI161002.1 GI:3694307

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
   Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 466)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jan 19, 1998 this sequence version replaced gi:2287379.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1599 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 454.

FEATURES Location/Qualifiers
source 1..466
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /map="839C02; 821G11; 4; 4p15.33-4p16.1; 4p15.33-4p16.1"
   /clone="IMAGE:1705398"
   /clone_lib="Soares_fetal_heart_NbHH19W"
   /sex="unknown"
   /dev_stage="19 weeks"
   /lab_host="DH10B (ampicillin resistant)"
   /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCAATATCTTTTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of

RESULT 5

LOCUS   AI185780   466 bp   mRNA   EST   29-OCT-1998
DEFINITION   qe44h04.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
    IMAGE:1741879 3', mRNA sequence.
ACCESSION   AI185780
NID   g3736418
VERSION   AI185780.1   GI:3736418
KEYWORDS   EST.
SOURCE   human.
   ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 466)
   AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
   TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
    Tumor Gene Index
   JOURNAL   Unpublished (1997)
COMMENT   On Jan 19, 1998 this sequence version replaced gi:2151491.

    Contact: Robert Strausberg, Ph.D.
    Tel: (301) 496-1550
    Email: Robert_Strausberg@nih.gov
    This clone is available royalty-free through LLNL ; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    Insert Length: 1622   Std Error: 0.00
    Seq primer: -40UP from Gibco
    High quality sequence stop: 466.

FEATURES   Location/Qualifiers
   source   1..466
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1741879"
    /clone_lib="Soares_fetal_lung_NbHL19W"
    /lab_host="DH10B (ampicillin resistant)"
    /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
    modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
    strand cDNA was primed with a Not I - oligo(dT) primer
    [5'-TGTTACCAATCTGAAGTGGGACGGCGCGCAATTTTTTTTTTTTTTTTTTT-3'],
    double-stranded cDNA was size selected, ligated to Eco RI
    adapters (Pharmacia), digested with Not I and cloned into
    the Not I and Eco RI sites of a modified pT7T3 vector
    (Pharmacia). Library went through one round of
    normalization to a Cot = 5. Library constructed by Bento
    Soares and M.Fatima Bonaldo. This library was constructed
    from the same fetus as the fetal lung library, Soares fetal
    lung NbHL19W."
BASE COUNT   79 a   183 c   123 g   80 t   1 others
ORIGIN

Query Match   1.9%;   Score 220.2; DB 43; Length 466;
Best Local Similarity   86.5%;   Pred. No. 1e-37;
Matches 243; Conservative   0; Mismatches 38; Indels   0; Gaps   0;

Qy   6330 acacagctgtaatcagccccaggagacacacctgggccccacccttctcatggaccctgcaagcta 6389
    ||||||||| |||| ||||| |||||||||||||| |||||||| |||||||||||||||| | |
Db   57 ACACAGCTGTGATCAGTCCCCAGGAGATCCCAGCGCTTCATCGGGCTCCTCCCGTGGCCA 116

Qy   6390 cctgctcctatacatggagacacaacctggggccacgcctgagggggctcctcactggaccttca 6449
    |||||||| | || | ||||||| |||||||| |||| ||||||||||||||||||||| |
Db   117 CCTGCTCAGTGCACGGAGACCACCAGGAGCCACCACCGGGCTCTCCATCGGCTCCTCCGGCCCTCA 176

Qy   6450 atggtcgccgccgccctgagtcgtccggcctcttaacactccacctggcccctgg 6509
    | ||| ||||||||||||||| |||||| |||||||| || ||||||||||| |||
Db   177 ACGGGCGCCGGCGGCTCGCCTGAGCTCGTCCCGTACTCAACGCCCTACTGGACCCTCA 236

Qy   6510 ccctggctaacctaatgtggtcccaggcaggcaagtcagcgcccac 6569
    |||||||||||||||||||||||||||||||||||||||| ||||||
Db   237 CCCTGGCCAACCTCAATGGGTCCCAGGCAGCGGTCGGGGGGACAAACCTCGTGTGCCACGCCC 296

Qy   6570 gagacggcagcattctggctggctgtcctgctctgtggc 6610
    ||||||||||||||||||||||||||||||||||||||||||
Db   297 GTGACGGCAGCAATCCTGGCTGCTGCTCCTCGCTCTATGTTGGC 337

RESULT 6

LOCUS   AI185924   477 bp   mRNA   EST   29-OCT-1998
DEFINITION   qe50c05.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
    IMAGE:1742408 3', similar to TR:Q16354 Q16354 PROLACTIN RECEPTOR ;,
    mRNA sequence.
ACCESSION   AI185924
NID   g3736562
VERSION   AI185924.1   GI:3736562
KEYWORDS   EST.
SOURCE   human.
   ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 477)
   AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
   TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
    Tumor Gene Index
   JOURNAL   Unpublished (1997)
COMMENT   On Feb 17, 1998 this sequence version replaced gi:2150926.

    Contact: Robert Strausberg, Ph.D.
    Tel: (301) 496-1550
    Email: Robert_Strausberg@nih.gov
    This clone is available royalty-free through LLNL ; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    Insert Length: 1552   Std Error: 0.00
    Seq primer: -40UP from Gibco
    High quality sequence stop: 467.

FEATURES   Location/Qualifiers
   source   1..477
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1742408"
    /clone_lib="Soares_fetal_lung_NbHL19W"
    /dev_stage="19 weeks"
BASE COUNT   79 a   183 c   123 g   80 t   1 others
ORIGIN

Query Match   1.9%;   Score 220.2; DB 43; Length 466;
Best Local Similarity   86.5%;   Pred. No. 1e-37;
Matches 243; Conservative   0; Mismatches 38; Indels   0; Gaps   0;

Qy   6330 acacagctgtaatcagccccaggagacacacctgggccccacccttctcatggaccctgcaagcta 6389
    ||||||||| |||| ||||| |||||||||||||| |||||||| |||||||||||||||| | |
Db   57 ACACAGCTGTGATCAGTCCCCAGGAGATCCCAGCGCTTCATCGGGCTCCTCCCGTGGCCA 116

Qy   6390 cctgctcctatacatggagacacaacctggggccacgcctgagggggctcctcactggaccttca 6449
    |||||||| | || | ||||||| |||||||| |||| ||||||||||||||||||||| |
Db   117 CCTGCTCAGTGCACGGAGACCACCAGGAGCCACCACCGGGCTCTCCATCGGCTCCTCCGGCCCTCA 176

Qy   6450 atggtcgccgccgccctgagtcgtccggcctcttaacactccacctggcccctgg 6509
    | ||| ||||||||||||||| |||||| |||||||| || ||||||||||| |||
Db   177 ACGGGCGCCGGCGGCTCGCCTGAGCTCGTCCCGTACTCAACGCCCTACTGGACCCTCA 236

Qy   6510 ccctggctaacctaatgtggtcccaggcaggcaagtcagcgcccac 6569
    |||||||||||||||||||||||||||||||||||||||| ||||||
Db   237 CCCTGGCCAACCTCAATGGGTCCCAGGCAGCGGTCGGGGGGACAAACCTCGTGTGCCACGCCC 296

Qy   6570 gagacggcagcattctggctggctgtcctgctctgtggc 6610
    ||||||||||||||||||||||||||||||||||||||||||
Db   297 GTGACGGCAGCAATCCTGGCTGCTGCTCCTCGCTCTATGTTGGC 337

RESULT 7

```
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCGAAGTGGGAGGCGGCCGCAATTTTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot - 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHH19W."
```

BASE COUNT   81 a   183 c   129 g   83 t   1 others
ORIGIN

Query Match   1.9%;   Score 220.2; DB 43;   Length 477;
Best Local Similarity   86.5%;   Pred. No. 1.1e-37;
Matches 243; Conservative   0; Mismatches 38; Indels   0; Gaps   0;

```
Qy  6330  acacagctgtaatcagcccccaggaccccacccttctcatcgtctccctgcaagcta 6389
          |||||||  ||||  ||||| |||||||||||||||||||||||||||||||| |||||
Db    57  ACACAGCTGTGATCAGTCCCCAGGATCCCACGCTTCTCATCCGGCTCCCTGCTGGCCA 116

Qy  6390  cctgctcatacatggagacacacacctgtgggggccaccgctgaggggggtctactggaccttca 6449
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   117  CCTGCTCAGTGACAGGAGACACCACCAGAGCCACCACCTGCCTTCTACTGGACCCTCA 176

Qy  6450  atggtcgtcgcgcctgagcgtgtccgcgtccctcctaacacctccacctgccctgg 6509
          | ||  ||||||||||| ||||||||||||||||||||||||||||||  ||||||
Db   177  ACGGGCGCGCGCCTGAGCTCTCCGCGTACTCAACGCCTTCAACCTTGGCTCTCTGG 236

Qy  6510  ccctggctaaccttaatgggtccaggagagcaagcaatctgttgttgtcacgccc 6569
          | |  ||||||||||||| ||  ||||||||| ||||||||||||| ||||  |||
Db   237  CCCTGGCCAACCTCAATGGGTCCAGGACAACCTCGTGTGTCCACGCC 296

Qy  6570  gagacggcaagcaattctggctggtcctgcctctctatgttggc 6610
          || ||||||||| |||||||||||||||||||||||||||||||||
Db   297  GTGACGGCAGCAGCATCCTGGCCTGGTCCTGCCTCTCTATGTTGGC 337
```

ACCESSION AI187074
LOCUS   AI187074   447 bp   mRNA   EST   29-OCT-1998
DEFINITION qe38a03.s1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
           IMAGE:1741228 3', mRNA sequence.
ACCESSION AI187074
NID   g3737712
VERSION   AI187074.1   GI:3737712
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 447)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1667   Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 431.
Location/Qualifiers
1..447
/organism="Homo sapiens"

RESULT 8

```
/db_xref="taxon:9606"
/clone="IMAGE:1741228"
/clone_lib="Soares_fetal_lung_NbHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCGAAGTGGGAGGCGGCCGCAATTTTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot - 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHH19W."
```

BASE COUNT   69 a   175 c   128 g   74 t   1 others
ORIGIN

Query Match   1.9%;   Score 220.2; DB 43;   Length 447;
Best Local Similarity   86.5%;   Pred. No. 1e-37;
Matches 243; Conservative   0; Mismatches 38; Indels   0; Gaps   0;

```
Qy  6330  acacagctgtaatcagcccccaggaccccacccttctcatcgtctccctgcaagcta 6389
          |||||||  ||||  ||||| |||||||||||||||||||||||||||||||| |||||
Db    75  ACACAGCTGTGATCAGTCCCCAGGATCCCACGCTTCTCATCCGGCTCCCTGCTGGCCA 134

Qy  6390  cctgctcatacatggagacacacacctgtgggggccaccgctgaggggggtctactggaccttca 6449
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   135  CCTGCTCAGTGACAGGAGACACCACCAGAGCCACCACCTGCCTTCTACTGGACCCTCA 194

Qy  6450  atggtcgtcgcgcctgagcgtgtccgcgtccctcctaacacctccacctgccctgg 6509
          | ||  ||||||||||| ||||||||||||||||||||||||||||||  ||||||
Db   195  ACGGGCGCGCGCCTGAGCTCTCCGCGTACTCAACGCCTTCAACCTTGGCTCTCTGG 254

Qy  6510  ccctggctaaccttaatgggtccaggagagcaagcaatctgttgttgtcacgccc 6569
          | |  ||||||||||| ||  ||||||||| ||||||||||||| ||||  |||
Db   255  CCCTGGCCAACCTCAATGGGTCCAGGACAACCTCGTGTGTCCACGCC 314

Qy  6570  gagacggcaagcaattctggctggtcctgcctctctatgttggc 6610
          || ||||||||| |||||||||||||||||||||||||||||||||
Db   315  GTGACGGCAGCAGCATCCTGGCCTGGTCCTGCCTCTCTATGTTGGC 355
```

RESULT   8
ACCESSION AI269388
LOCUS   AI269388   417 bp   mRNA   EST   27-JAN-1999
DEFINITION qi26b05.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1857585
           3', mRNA sequence.
ACCESSION AI269388
NID   g3888555
VERSION   AI269388.1   GI:3888555
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1634   Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 406.

```
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1857585"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT         64 a      164 c      116 g       73 t
ORIGIN
```

Query Match   1.9%;   Score 220.2;   DB 44;   Length 417;
Best Local Similarity 86.5%;   Pred. No. 1e-37;
Matches 243; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
Qy  6330  acacaagctgtaatcaagccccaggagaccccaccaccccttctctatcgcgtctccctgcaagcta 6389
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||
Db    75  ACACAGCTGTGATCAGTCCCCAGGAATCCCACGCGTTCTCATCGGCTGCTCCCCTGCTGGCCA  134

Qy  6390  cctgctcctatacatggagacacacctggggccacggccctaactggaacttca 6449
          ||||| || ||||||||||| ||||||| ||||| ||||||||| |||||| |||
Db   135  CCTGCTCAGTGCACGGAGACACCACCAGGAGCCACCGCCGGAGGCCTTCTCAGCTA  194

Qy  6450  atggtcgccgcccctgcgctcctcgagctcttaacacctccaaccctggcccctgg 6509
          || ||||||||||||| || || |||| || |||||||| ||||| |||||
Db   195  ACGGGCGCGGGCGCCCTGCCCGTGTACTCAAACGCCTCCACCTTGGCTGCTGG  254

Qy  6510  ccctggctaaacttaatgggtcaggcagcagtcaaggagacaatctgtgttgtcacgccc 6569
          |||||||||||||| |||||||| ||||| ||||| |||||||| ||||||||||||
Db   255  CCCTGGCCCAACCTCAATGGGTCAGGCAGCGGTCGGGGACAACCTCGTGTGCCACGCCC  314

Qy  6570  gagacggcagcaatcctggctgtggcccctgccctcatgttggc 6610
          ||||||||| || |||||||||||||||||||||||||||||||
Db   315  GTGACGGCAGCAATCCTGGCTGGCTTCCGTACTTTGGC  355
```

```
RESULT  9
LOCUS      AI579568        332 bp    mRNA                      EST       05-APR-1999
DEFINITION UI-R-GO-ut-h-08-0-UI s3 UI-R-GO Rattus norvegicus cDNA clone
           UI-R-GO-ut-h-08-0-UI 3', mRNA sequence.
ACCESSION  AI579568
NID        g4563944
VERSION    AI579568.1  GI:4563944
KEYWORDS   EST.
SOURCE     Norway rat.
  ORGANISM Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE  1  (bases 1 to 332)
  AUTHORS  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE    Normalization and subtraction: two approaches to facilitate gene
           discovery
  JOURNAL  Genome Res. 6 (9), 791-806 (1996)
  MEDLINE  97044477
COMMENT    On Mar 10, 1998 this sequence version replaced gi:2948398.
```

University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu

Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward.

```
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="Rattus norvegicus"
                     /strain="sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-GO-ut-h-08-0-UI"
                     /clone_lib="UI-R-GO"
                     /dev_stage="adult"
                     /lab_host="DH10B (life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-GO
                     library is a normalized library constructed from a
                     mixture of rat tissues (nodose ganglia, dorsal root
                     ganglia, and trigeminal ganglia). The tag is a string of
                     6 nucleotides present between the Not I site and the
                     oligo-dT track. The library was constructed as described
                     by Bonaldo, Lennon and Soares, Genome Research 6:
                     791-806, 1996."
BASE COUNT         49 a      121 c       93 g       69 t
ORIGIN
```

Query Match   1.9%;   Score 219.6;   DB 48;   Length 332;
Best Local Similarity 92.4%;   Pred. No. 1.3e-37;
Matches 231; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy  6330  acacaagctgtaatcaagccccaggagaccccaccaccccttctctatcgcgtctccctgcaagcta 6389
          ||||||||| ||| |||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    83  ACACAGCTGTGATCAGTCCCCAGGAGACCCCACCACTCTTCTCATCGGATCCTCCCTTCAGCTA  142

Qy  6390  cctgctcctatacatggagacacacctggggccacccgctgagggggctctactggaccttca 6449
          |||||||||||| ||||||||||||||||||||||||||||||||| |||||||||||||||
Db   143  CGTGCTCTATACATGGAGACACACCGGGGCCACTGCTGAGGGCCTTCTACTGGACCCTCA  202

Qy  6450  atggtcgccgcccctgcgctcctccgccctccttaacacctccaacccctggcccctgg 6509
          ||||||||| |||||||||| |||||||| |||||||||||||||||||||||||||||
Db   203  ACGGCCGCGGCCCCTGCGCTCCAGAGCTGTCCGGTCTCTCAACACCTCCACCCTGGCCTTGG  262

Qy  6510  ccctggctaaacctaatgggtcggcagtcagtcaggaggaacaatctggtgttgtcacgccc 6569
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263  CCCTGGCTAACCTTAATGGGTCCAGCAGCAGCAGTCAGGGACAATCCTGGTGGTGTCAGGCCC  322

Qy  6570  gagacggcag 6579
          |||| |||||
Db   323  GAGATGGCAG  332
```

```
RESULT  10
LOCUS      AI333812        431 bp    mRNA                      EST       13-FEB-1999
DEFINITION qp93e12.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
           IMAGE:1930606 3' similar to TR:Q16354 Q16354 PROLACTIN RECEPTOR ;;
           mRNA sequence.
ACCESSION  AI333812
NID        g4070371
VERSION    AI333812.1  GI:4070371
KEYWORDS   EST.
SOURCE     human.
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 431)
  AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
Best Local Similarity  93.5%;  Pred. No. 3.2e-36;
Matches 244;  Conservative  0;  Mismatches 15;  Indels  2;  Gaps  2;

Qy    11533  caggtcctgccggctaaactctaaggataggccatcctcct-gctgggtcagacctggag  11592
             |||||||||||||||||||||| |||||||||||||| |||||||||| |||||||||||
Db        2  CACGTTCTGCGGGGTTAACTATTATAGGATCGGCTATCCTCTTCGCGTGGGTTAGACCTGGAG  61

Qy    11592  gctcacctgaattggaccccctctgtaccactctgggcaacaaagaaacctaccagaggct  11651
             |||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db       62  GCTCACCTGAATTGGAGCCCCTCTTGTACCACATTTGGCAACAAAAGAAACCTACCAGAGGCT  121

Qy    11652  -ggggcacaatgagctccccacaaaccacacagcttggtccacatggtcacacttggatat  11710
             |||||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db      122  GGGGCACAATGAGCTCCCACAACCACACAGCTTTGGTCCACATGGTCACACTTGGATAT  181

Qy    11711  accccagtgtgggtaggggttgggggtattgcaagggccctcccaagagagtcttttaaa  11770
             |||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db      182  ACCCCAGTGTGGGTAGGGGTTGGGGGTATTGCAAGGGCCCTCCCAAGAGAGTCTTTTAAATAAATAAAT  241

Qy    11771  aaaggaggttgttcaggtcccg  11791
             ||||||||||||||||||||||
Db      242  AAAGGAGGTTGTTCAGGTCCCG  262
```

```
RESULT 14
LOCUS        AI574687        319 bp        mRNA        EST        30-MAR-1999
DEFINITION   UI-R-G0-uc-b-11-0-UI.s1 UI-R-G0 Rattus norvegicus cDNA clone
             UI-R-G0-uc-b-11-0-UI 3', mRNA sequence.
ACCESSION    AI574687
NID          g4544683
VERSION      AI574687.1  GI:4544683
KEYWORDS     EST.
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE    1  (bases 1 to 319)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
             discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      On May 18, 1998 this sequence version replaced gi:3133307.
             Contact: Soares, MB
             Program for Rat Gene Discovery and Mapping
             University of Iowa
             451 Eckstein Medical Research Building Iowa City, IA 52242, USA
             Tel: 319 335 8250
             Fax: 319 335 9565
             Email: msoares@blue.weeg.uiowa.edu
             Oligo-dT track not found, Not I site shown in beginning of sequence
             is likely internal to the message. cDNA Library Preparation: M.B.
             Soares Lab Clone distribution: clones will be available through
             Research Genetics (www.resgen.com)
             Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1..319
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /map="6p21.3; 5q33"
                     /clone="UI-R-G0-uc-b-11-0-UI"
                     /clone_lib="UI-R-G0"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT71D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-G0
                     library is a normalized library constructed from a
                     mixture of rat tissues (nodose ganglia, dorsal root
                     ganglia, and trigeminal ganglia). The tag is a string of
                     6 nucleotides present between the Not I site and the
```

```
BASE COUNT     46 a     116 c     88 g     68 t     1 others
ORIGIN         oligo-dT track. The library was constructed as described
               by Bonaldo, Lennon and Soares, Genome Research 6:
               791-806, 1996."

Query Match           1.8%;  Score 207.2; DB 48;  Length 319;
Best Local Similarity 92.0%;  Pred. No. 6.2e-35;
Matches 218;  Conservative  0;  Mismatches 19;  Indels  0;  Gaps  0;

Qy    6330  acacagctgtaatcagccccccaggagaccccaccacccttcctcatcggctgctcccctccaagcta  6389
            ||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db      83  ACACAGCTGTGATCAGTCCCCAGGACCCCAGGAGACCCCACTCTTCTCATCGGATCCTCCCTTCATGCTA  142

Qy    6390  cctgctctatacatggagacacacacctggggccaccgctgagggggctactggaccttca  6449
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     143  CGTGCTCTATACAGGGAGACACACCGGGGGCCACTGCTGCAGGGGCCTCTTACTGGACCCTCA  202

Qy    6450  atggtccgcccctgacccctgagcctgtcccgcctccttaaacacctcacacctggcctgg  6509
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     203  ACGGCCGCGGCCGCTGCCCCTTGCACAGAGCTGTCNCGCGTCCTCCNACACCTCCACACCTGGCCCCTTGG  262

Qy    6510  ccctggcctaaccttaattgggtccaggcagcagtcaggagacaatctggtgttgtcacg  6566
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     263  CCCTGGCTGAACCTTAATGGGTCCAGGCAGCAGTCAGGGGCAGTCAGGGGGCAGCAAATCTGGTGTGTCACG  319
```

```
RESULT 15
LOCUS        W17583        390 bp        mRNA        EST        10-SEP-1996
DEFINITION   mb75b01.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
             IMAGE:335209 5', mRNA sequence.
ACCESSION    W17583
NID          g1291995
VERSION      W17583.1  GI:1291995
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 390)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubque,T.,
             Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,T.,
             Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
             Thelsing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
             Waterston,R.
TITLE        The WashU-HHMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      On May 9, 1995 this sequence version replaced gi:804170.
             Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:216609
             Seq primer: mob.REGA+ET
             High quality sequence stop: 384.
FEATURES             Location/Qualifiers
     source          1..390
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="3"
                     /clone="IMAGE:335209"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT71D3 (Pharmacia) with a modified
```

ORIGIN
BASE COUNT   89 a   111 c   114 g   76 t

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGGGGGCCGCATTTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

Query Match   1.7%;   Score 200.2; DB 26; Length 390;
Best Local Similarity   97.4%;   Pred. No. 2.2e-33;
Matches 225;   Conservative 0;   Mismatches 3; Indels 3; Gaps 2;

Qy 11534 aggtcctgccggctaaactctaaggataggccatcctctgctgggtcagacctggaggc 11593
           |||||||||||||||| |||||||||| || |||||||||||||||||| ||||||||||
Db   160   AGGTCCTGCCGGCTAAACTCTAAGGATAGGCCATCCTCTGCTGGGTCAGACCTGGAGGC 219

Qy 11594 tcacctgaattggagcccctcctgtaccatctgggcaacaaagaaacctaccagaggct-g 11652
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   220   TCACCTGAATTGGAGCCCCTCTGTACCATCTGGGCAACAAAGAAACCTACCAGAGGCTGG 279

Qy 11653 ggcacaatgagctcccacaaccagcttttggtccacatgatggtcacacttggatatac 11712
           |||||||||||||||||||||||||||| ||||| |||||||||||||||||| |||||
Db   280   GGCACAATGAGCTCCCACAACCACAGCTTTGGTCCACATGATGGTCACACTTGGATATAC 339

Qy 11713 cccagtg--tgggtaggtgtttgggtattgcaggggccctcccaagagtctctt 11761
           ||||||| || |||| |||||||||||||||||||||||||||||||||||||
Db   340   CCCAGTGGGTGGGTAAGGGTTGGGGTATTCCAGGGGCCCTCCCAAGAGTCTCTT 390

Search completed: September 19, 1999, 15:16:34
Job time: 13168 sec

First Record

TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Jan 14, 1998 this sequence version replaced gi:1797892.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1615   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 407.
FEATURES   Location/Qualifiers
     source   1..431
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="4p16.1-4pter"
                     /clone="IMAGE:1930606"
                     /clone_lib="Soares_fetal_lung_NbHl19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGACGGCGCCGCCAATTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library. Soares
                     fetal heart NbHH19W."
BASE COUNT   68 a   169 c   122 g   72 t
ORIGIN

RESULT 11
Query Match   86.4%;   1.9%; Score 219.2; DB 45; Length 431;
Best Local Similarity
Matches 242; Conservative   0; Mismatches 38; Indels 0; Gaps 0;

Qy   6331 cacagctgtaatcagcccccaggacccccacccttctcatcggctcctccctgcaagctac 6390
           |||||||| |||||| ||||| ||||||||||||||||||| |||||||||||||| ||
Db   58 CACAGCTGTGATCAGTCNCCCACGGGGATCCCACGCTTCTCATCGGCTCCTCCCTGGCCAC 117

Qy   6391 ctgctcttatacatggagacacacctgggggccacgcgctgaggggtctcactggacctcaa 6450
           |||||| || || ||||| ||| ||| |||| ||| ||| |||||| |||||||| |||
Db   118 CTGCTCAGTGCACGGAGACCCACCAGGAGCCCGAGGGCCTTACTGGACCCTCAA 177

Qy   6451 tggtcgcgcgccctgccctctgagcgtgtcccgcctccttaacacctccaccctggcctggc 6510
           ||||| ||||||||||||| || |||| | ||||||||||||| ||||||| ||||| |||
Db   178 CGGGGCCGCGCCCTGCCCTCCCGTGTACTCAACGGCCTCCACCTTGGCTCTGGC 237

Qy   6511 cctggctaacttaatgggtccaggaggcagtcaggaggagacaatcctcggtcgcctcacgccg 6570
           ||||| |||| |||||||||||||||| || ||| ||| ||||||||||| || ||||| |||
Db   238 CCTGGCCAACCTCAATGGGTCCAGGAGGCGGTTCGGGGGGACAACCTCGTGTGCCACGCCCG 297

Qy   6571 agacggcagcagcattctggctgcctcctcgcctctatgttggc 6610
           | |||||| | ||||||||||||| |||| |||| ||||||||||
Db   298 TGACGGCGACATCCTGGCTCTGCCTCTCGCCTTATGTTGGC 337

LOCUS   AI670108   385 bp   mRNA   EST   14-MAY-1999
DEFINITION we65f03.x1 Soares_thymus_NHFTh Homo sapiens cDNA clone
            IMAGE:2345981 3' similar to TR:075462 075462 CYTOKINE-LIKE FACTOR-1
            PRECURSOR.;, mRNA sequence.
ACCESSION   AI670108
NID   g4834882
VERSION   AI670108.1 GI:4834882
KEYWORDS   EST.

Second Record

SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 385)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   On May 9, 1996 this sequence version replaced gi:1133091.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 372.
FEATURES   Location/Qualifiers
     source   1..385
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="2"
                     /clone="IMAGE:2345981"
                     /clone_lib="Soares_thymus_NHFTh"
                     /dev_stage="fetal"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: thymus; pooled; Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'-
                     TGTTACCAATCTGAAGTGGGACGGCGCCGCCAACGTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adapters
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   55 a   156 c   108 g   66 t
ORIGIN

RESULT 12
Query Match   85.4%;   1.8%; Score 215.4; DB 50; Length 385;
Best Local Similarity
Matches 240; Conservative   0; Mismatches 41; Indels 0; Gaps 0;

Qy   6330 acacagctgtaatcagcccccaggacccccacccttctcatcggctcctccctgcaagcta 6389
           ||||||||| |||||| ||||| ||||||||||||||||| |||||||||||||| ||
Db   75 ACACAGCTGTGATCAGTCNCCCAGGATCCCACGCTTCTCATCGGCTCCTCCCTGGCCA 134

Qy   6390 cctgctcttatacatggagacacacctgggggccacgcgctgaggggtctcactggacctcaa 6449
           |||||| || || ||||| ||| ||| |||| ||| ||| |||||| |||||||| |||
Db   135 CCTGCTCAGTGCACGGAGACCCACCAGGAGCCCGAGGGCCTTACTGGACCCTCA 194

Qy   6450 atggtcgcgcgccctgccctctgagcgtgtcccgcctccttaacacctccaccctggcctgg 6509
           ||||| ||||||||||||| || |||| | ||||||||||||| ||||||| ||||| |||
Db   195 ACGGGGCCGCGCCCTGCCCTCCCGTGTACTCAACGGCCTCCACCTTGGCTCTGG 254

Qy   6510 cctggctaacttaatgggtccaggaggcagtcaggaggagacaatcctcggtcgcctcacgccg 6569
           ||||| |||| |||||||||||||||| || ||| ||| ||||||||||| || ||||| |||
Db   255 CCCTGGCCAACCTCAATGGGTCCAGGAGGCGGTTCGGGGGGACAACCTCGTGTGCCACGCCC 314

Qy   6570 gagacggcagcagcattctggctgcctcctcgcctctatgttggc 6610
           | |||||| | ||||||||||||| |||| |||| ||||||||||
Db   315 GTGACGGCGACATCCTGGCTCTTGCCTCTCGCCTTATGTTGGC 355

LOCUS   W66776   482 bp   mRNA   EST   14-JUN-1996
DEFINITION W66776 me17bl1.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:387741 5' similar to PIR:B38252 B38252 granulocyte
            colony-stimulating factor receptor precursor ;, mRNA sequence.
ACCESSION   W66776

sequence comparison (handwritten annotation)

S

```
NID        g1375694
VERSION    W66776.1  GI:1375694
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1  (bases 1 to 482)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
           Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
           Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
           Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
           Waterston,R.
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    On Apr 14, 1993 this sequence version replaced gi:785250.

           Contact: Marra M/Mouse EST Project
           WashU-HHMI Mouse EST Project
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:23573
           Possible reversed clone: similarity on wrong strand
           Seq primer: EPrimer
           High quality sequence stop: 359.

FEATURES   Location/Qualifiers
source     1..482
           /organism="Mus musculus"
           /strain="C57BL/6J"
           /db_xref="taxon:10090"
           /map="10 pter-cen"
           /clone="IMAGE:387741"
           /clone_lib="Soares mouse embryo NbME13.5 14.5"
           /sex="unknown"
           /tissue_type="embryo"
           /dev_stage="13.5-14.5dpc total fetus"
           /lab_host="DH10B"
           /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
           polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
           was primed with a Not I - oligo(dT) primer [5'
           TGTTACCAATCTGAAGTGGGACGGCGCGGGAAATTTTTTTTTTTTTTTTTTTTT
           T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
           14.5pc embryos [total RNA provided by Minoru Ko, Wayne
           State Univ., from 2 ]; double-stranded cDNA was ligated to
           Eco RI adaptors (Pharmacia), digested with Not I and
           cloned into the Not I and Eco RI sites of the modified
           pT7T3 vector. Library went through one round of
           normalization, and was constructed by Bento Soares and
           M.Fatima Bonaldo."

BASE COUNT    96 a    138 c    147 g    101 t
ORIGIN
```

Sequence (Comparison) R.

```
Query Match      1.8%;    Score 214.8;   DB 26;   Length 482;
Best Local Similarity   81.1%;
Matches 305;  Conservative   0;  Mismatches  2;  Indels  69;  Gaps  2;

Qy  8604  gtgacaacgacccccaccccgacgtgcacgtgagccgcgttggggggcctggaggaccag  8663
          |||||||| |||||||  |  |||||||||||||||||||   |||||||||||||||||
Db   317  GTGACAACGGACCCCCACCCACAGCGTGCACGTGAGCCGCGGCCTGGGGGGCCTGGAGGACCAG  258

Qy  8664  ctgagtgtgcgctgggtctctcacacacacagctctccaaggattccctcttccaagccaagtac  8723
          |||||||||| |||||||||| |  |  |  ||||| |||||||||||||||||||||||||||||
Db   257  CTGAGTGTGCTGCTGGGTCTCACCACCACCAGCTCTTCAAGGATTCCCTCTTCCAAGCCAAGTAC  198

Qy  8724  cag-atccgctaccgcgtggaggacagcgtggactggaagtgcgcccgtcccgccccggac  8782
          ||| |||||||||||||||||||||| ||||||||||
Db   197  CAGTATTCGCTACCGCGGTGGAGGACAACGGTGGACTGGA--------------------  159
```

```
Qy  8783  ccgcccctgacccgccccgccccgcaatctgactcctccctcaccgtgcaggtggtggtggatgac  8842
                                                            ||||||||||  |||
Db   159  ------------------------------------------------AGGTGGTGGGGATGAC  146

Qy  8843  gtcagcaaccaggacctcctgcgtctcgcgggcctgaagcacggcgacgttacttcgtc  8902
          ||||||||||||||||||||||||||||||||||||||||  ||||| | |||||||||
Db   145  GTCAGCAACCAGGACCTCCTGCGTCTCGCGGGCCTGAAGCCCGGCGCACCGTTACTTCGTC  86

Qy  8903  caagtggcggttgtaaccccattcgggatctatggtcggaaaaaggcgcggagaatctggagcgag  8962
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    85  CAAGTGGCGGTTGTAACCCCATTCGGGATCTATGGGTCGGAAAAAGGCGCGGAGAATCTGGAGCGAG  26

Qy  8963  tggagcgcaccccaccg  8978
          ||||||||||||| |||
Db    25  TGGAGCGCACCCCATCG  10
```

```
RESULT 13
AV032198
LOCUS      AV032198    262 bp    mRNA         EST       03-JUN-1999
DEFINITION AV032198 Mus musculus adult C57BL/6J cerebellum Mus musculus cDNA
           clone 1500034F20, mRNA sequence.
ACCESSION  AV032198
NID        g4831746
VERSION    AV032198.1  GI:4831746
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1  (bases 1 to 262)
AUTHORS    Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Ito,M., Aizawa,K.,
           Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
           Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
           Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
           Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
           Sugahara,Y., Suzuki,J., Suzuki,H., Tateno,M., Tomaru,Y.,
           Tominaga,N., Watanabe,S., Yagame,M., Yamanaka,T., Yokota,T.,
           Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
TITLE      RIKEN Mouse ESTs
JOURNAL    Unpublished (1999)
COMMENT    On Jun 22, 1998 this sequence version replaced gi:3246782.

           Contact: Chie Owa
           Genome Science Laboratory
           RIKEN
           3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
           Tel: 81-298-36-9145
           Fax: 81-298-36-9098
           Email: genome-res@rtc.riken.go.jp
           Thermostabilization and thermoactivation of thermolabile enzymes by
           trehalose and its application for the synthesis of full length cDNA
           (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
           Transcriptional sequencing: A method for DNA sequencing using RNA
           polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998)) for
           Please visit our web site (http://genome.rtc.riken.go.jp) for
           further details.

FEATURES   Location/Qualifiers
source     1..262
           /organism="Mus musculus"
           /strain="C57BL/6J"
           /db_xref="taxon:10090"
           /clone="1500034F20"
           /clone_lib="Mus musculus adult C57BL/6J cerebellum"
           /sex="male"
           /tissue_type="cerebellum"
           /dev_stage="adult"

BASE COUNT     63 a     63 c     69 g     67 t
ORIGIN

Query Match      1.8%;    Score 213;   DB 50;   Length 262;
```

PR   03-FEB-1989; US-306503.
PA   (GENE-) GENETICS INST INC.
PA   (WHIT-) WHITEHEAD INST.
PI   D'andrea A, Wong G;
DR   WPI; 90-260931/34.
DR   N-PSDB; Q05748.
PT   Erythropoietin receptor and gene - used for developing reagents
PT   and systems to control and study erythropoiesis.
PS   Disclosure; Fig 2; 53pp; English.
CC   The sequence was deduced from DNA obtd. from a clone isolated from
CC   a commercially available human genomic cDNA library in phage
CC   lambda Fix (Stratagene). The sequence encodes a type I trans-
CC   membrane protein with binding affinity for EPO. The gene and
CC   recombinant EPO receptor produced on expression of the gene are
CC   used to develop reagents and systems to control and study
CC   erythropoiesis. It is believed that the EPO receptor is dys-
CC   functional in individuals with Diamond Blackfan anaemia, and may
CC   be hyperactive in polycythemia vera.
CC   See also R0511 (murine EPO receptor).
SQ   Sequence 508 AA;

Qy   1 WSXWS 5.
        || |||
Db   233 WSAWS 237

Query Match   96.8%;   Score 30;   DB 1;   Length 508;
Best Local Similarity   80.0%;   Pred. No. 4.8e+02;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT   3
R24017
ID   R24017 standard; Protein; 552 AA.
AC   R24017;
DT   26-NOV-1992 (first entry)
DE   Fusion protein GM-CSFRFc.
KW   Granulocyte macrophage-colony stimulating factor; GM-CSF; IgG1;
KW   Immunoglobulin G1.
OS   Synthetic.
FH   Key   Location/Qualifier
FT   region   1..318
FT   /note= "human GM-CSF receptor"
FT   region   319..336
FT   /note= "linker and hinge"
FT   region   337..445
FT   /note= "IgG1 CH2"
FT   region   446..552
FT   /note= "IgG1 CH3"
PN   EP-488170-A.
PD   03-JUN-1992.
PR   26-NOV-1991; 120187.
PR   28-NOV-1990; DE-037837.
PA   (BEHW) BEHRINGWERKE AG.
PI   Lauffer L, Oquendo P, Zettlmeissl G;
DR   WPI; 92-185084/23.
PT   Cell free receptor binding test contg. recombinant fusion protein -
PT   comprising carrier bound to fusion partner coupled to fixed support,
PT   and second, labelled binding partner, for receptor or antibody
PT   screening etc.
PS   Example; Fig 12; 24pp; German.
CC   The sequence is that of fusion protein GM-CSFRFc comprising the
CC   extracellular domain of granulocyte macrophage-colony stimulating
CC   factor (GM-CSF) fused via a hinge region to the Fc part of the heavy
CC   chain of human IgG1. It may be used as part of a cell free receptor
CC   binding test which can be used for the identification of agonists,
CC   antagonists, antibodies, biological activity of soluble cellular
CC   receptors, functional substances, functional analysis of modified ligands and diagnostic or
CC   therapeutic substances. See also R24016.
SQ   Sequence 552 AA;

Query Match   96.8%;   Score 30;   DB 1;   Length 552;

Best Local Similarity   80.0%;   Pred. No. 4.8e+02;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 WSXWS 5.
        || |||
Db   306 WSSWS 310

RESULT   4
R10795
ID   R10795 standard; Protein; 622 AA.
AC   R10795;
DT   25-APR-1991 (first entry)
DE   Human prolactin receptor.
KW   Human prolactin receptor; PRL; assay; antibody; growth hormone.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   peptide   1..24
FT   /label= sig_peptide
FT   protein   25..622
FT   /label= mat_protein
PN   US4992378-A.
PD   12-FEB-1991.
PR   16-DEC-1988; 286445.
PR   16-DEC-1988; US-286445.
PA   (ROYA-) ROYAL INST ADVAN LE.
PI   Kelly PA, Diane J;
DR   WPI; 91-065341/09.
DR   N-PSDB; Q10550.
PT   Isolated cDNA sequence encoding human prolactin receptor - useful
PT   for expressing the receptor, e.g. for screening assays and antibody
PT   prodn.
PS   Disclosure; Fig. 1 (A-E); 11pp; English.
CC   The human PRL receptor cDNA is isolated by screening a lambda gt 10
CC   library prepared from normal human hepatoma Hep G2 and T47-D breast
CC   cancer cells. Initially, 1x10 6 recombinants are screened with the
CC   following probe: a complementary RNA of the rat
CC   prolactin receptor. One positive recombinant was isolated from the
CC   Hep G2 library, which was later used as a probe to rescreen the
CC   library. Five additional cDNAs were identified, and by combining
CC   two of the six cDNAs, a PRL receptor cDNA of 2.5 kb was constructed,
CC   contg. a single ORF of 1866 bp. Similar partial length cDNA were
CC   isolated from the T47-D library.
CC   Several regions of sequence identity between the human growth
CC   hormone and PRL receptors can be found, both in the extracellular
CC   and cytoplasmic domains.
CC   The sequence is an important genetic engineering tool which may be
CC   used for the screening of growth hormone variants, for the development
CC   of test kits to measure PRL receptor levels in human breast and prostate
CC   cancer biopsies, for the measurement of bioactive forms of prolactin,
CC   and for the development of drugs to induce stimulation or inhibition of
CC   the immune system.
SQ   Sequence 622 AA;

Qy   1 WSXWS 5.
        || |||
Db   215 WSAWS 219

Query Match   96.8%;   Score 30;   DB 1;   Length 622;
Best Local Similarity   80.0%;   Pred. No. 5.4e+02;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT   5
R10919
ID   R10919 standard; Protein; 400 AA.
AC   R10919;
DT   08-MAY-1991 (first entry)
DE   Human GM-CSF receptor.
KW   Granulocyte-macrophage colony-stimulating factor; myeloid leukaemia;
KW   autoimmune disease.
FH   Key   Location/Qualifiers

OM protein - protein search, using sw model

Run on: September 17, 1999, 03:03:09 ; Search time 64.1 Seconds

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

(without alignments)
1.848 Million cell updates/sec

Title: US-09-037-657-1
Perfect score: 31
Sequence: 1 WSXWS 5

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	96.8	507	1	R06511	EPO receptor seque
2	30	96.8	508	1	R06512	EPO receptor, Eryt
3	30	96.8	552	1	R24017	Fusion protein GM-
4	30	96.8	622	1	R10795	Human prolactin re
5	30	96.8	400	1	R10919	Human GM-CSF recep
6	30	96.8	211	1	R22228	Truncated human pr
7	30	96.8	211	1	R24273	Truncated human pr
8	30	96.8	284	1	R23971	MPlV env-vmpl fusi
9	30	96.8	635	1	R23970	MPlV env protein w
10	30	96.8	382	1	R27659	MPlV-env related p
11	30	96.8	184	1	R27660	Partial MPlV-env r
12	30	96.8	507	1	R47517	MEL EPO receptor.
13	30	96.8	508	1	R47518	Human EPO receptor
14	30	96.8	265	1	R50326	Mouse soluble EPO
15	30	96.8	507	1	R50327	Mouse soluble EPO
16	30	96.8	507	1	R69502	Mouse erythropoiet
17	30	96.8	508	1	R69503	Human erythropoiet
18	30	96.8	5	1	R69782	Thrombospondin-der
19	30	96.8	508	1	R70032	Human erythropoiet
20	30	96.8	626	1	R75939	Murine myeloprolif
21	30	96.8	635	1	R75940	Human myeloprolife
22	30	96.8	482	1	R75941	Soluble murine MPL
23	30	96.8	633	1	R75908	Type I MPL recepto
24	30	96.8	633	1	R79053	Mouse type I MPL r
25	30	96.8	967	1	R83441	Mouse type I MPL r
26	30	96.8	423	1	R92814	Human interleukin-
27	30	96.8	432	1	R92813	Murine interleukin
28	30	96.8	972	1	R92288	Cycloisomaltooligo
29	30	96.8	934	1	R92289	Cycloisomaltooligo
30	30	96.8	32	1	R98939	Synthetic human er
31	30	96.8	422	1	R99090	Human interleukin-
32	30	96.8	441	1	R99091	Murine Etl-2 gene
33	30	96.8	633	1	W03513	Mouse type I MPL r
34	30	96.8	633	1	W09948	Mouse type I MPL r
35	30	96.8	488	1	W08349	EpoRFc fusion prot
36	30	96.8	528	1	W69963	Aeromonas caviae x
37	30	96.8	898	1	W78898	Rat UNC-5 homologu
38	30	96.8	1522	1	W99302	Human BAI2 protein
39	30	96.8	1584	1	W99300	Human BAI1 protein
40	29	93.5	323	1	P90527	B cell stimulating
41	29	93.5	344	1	P90528	B cell stimulating
42	29	93.5	386	1	P90526	B cell stimulating
43	29	93.5	468	1	P90525	B cell stimulating

ALIGNMENTS

| | 44 | 29 | 93.5 | 810 | 1 | R04574 | Derived amino acid |
| | 45 | 29 | 93.5 | 218 | 1 | W40287 | Human TSP1 protein |

RESULT 1
R06511
R06511; R06511 standard; protein; 507 AA.
ID
AC R06511;
DT 04-JAN-1991 (first entry)
DE EPO receptor sequence deduced from DNA of clone 190.
KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.
OS Mus musculus.
FH Key   Location/Qualifiers
FT peptide   1..24
FT   /label=signal peptide
FT domain   25..248
FT   /label=extracellular domain
FT   /note=EPO binding region
FT domain   248..271
FT   /label=transmembrane domain
FT domain   272..507
FT   /label=intracellular domain
FT modified_site   75..77
FT   /label=N-linked_glycos
FT modified_site   182..184
FT   /label=N-linked_glycos
PN WO9008822-A.
PD 09-AUG-1990.
PF 01-FEB-1990; U00635.
PR 03-FEB-1989; US-306503.
PA (WHIT-) WHITEHEAD INST.
PA (GENE-) GENETICS INST INC.
DR D'andrea A; Wong G;
DR WPI; 90-260931/34.
DR N-PSDB; C05747.
PT Erythropoietin receptor and gene - used for developing reagents
PT and systems to control and study erythropoiesis.
PS Disclosure; Fig 1; 53pp; English.
CC The sequence was deduced from DNA from a clone isolated from a
CC cDNA library prepd. from uninduced murine erythroleukemia cells.
CC It is a type I transmembrane protein with binding affinity for EPO.
CC The gene and recombinant EPO receptor produced on expression of
CC the DNA are used to develop reagents and systems to control and
CC study erythropoiesis. It is believed that the EPO receptor is
CC dysfunctional in individuals with Diamond Blackfan anaemia, and
CC may be hyperactive in polycythemia vera.
CC See also R06512 (human EPO receptor).
SQ Sequence 507 AA;

QY   1 WSXWS 5
    ||||| ||
Db   232 WSAWS 236

Query Match   96.8%;   Score 30;   DB 1;   Length 507;
Best Local Similarity   80.0%;   Pred. No. 4.4e+02;
Matches   4; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

RESULT 2
R06512
ID R06512; R06512 standard; protein; 508 AA.
AC R06512;
DT 04-JAN-1991 (first entry)
DE EPO receptor.
KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.
OS Homo sapiens.
PN WO9008822-A.
PD 09-AUG-1990.
PF 01-FEB-1990; U00635.

Sequence "A" Comparison (handwritten annotation)